

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2006, 05:43:07 ; Search time 51 Seconds

(without alignments)
919.930 Million cell updates/sec

Title: US-10-823-506-8

Perfect score: 2836
Sequence: 1 MAGAMTPPPVQPARPGF.....LPAKGEVQNALNDHGHRRH 536

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 650591

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: Issued Patents AA:*
2: /EMC_Celerra_sids3/prodata/2/iaa/5 COMB pep:*
3: /EMC_Celerra_sids3/prodata/2/iaa/6 COMB pep:*
4: /EMC_Celerra_sids3/prodata/2/iaa/7 COMB pep:*
5: /EMC_Celerra_sids3/prodata/2/iaa/H COMB pep:*
6: /EMC_Celerra_sids3/prodata/2/iaa/RE COMB pep:*
7: /EMC_Celerra_sids3/prodata/2/iaa/backfillseq1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2836	100.0	536	2 US-09-359-167-8	Sequence 8, Appli
2	2804.5	98.9	585	2 US-09-949-016-7705	Sequence 7705, Ap
3	2621	92.4	495	2 US-09-359-167-2	Sequence 2, Appli
4	2621	92.4	495	2 US-09-915-181A-7	Sequence 7, Appli
5	2329	82.1	495	2 US-09-359-167-4	Sequence 4, Appli
6	1698	59.9	495	2 US-09-359-167-10	Sequence 10, Appli
7	1698	59.9	495	2 US-09-359-167-12	Sequence 12, Appli
8	970	34.2	582	2 US-09-740-041-4	Sequence 4, Appli
9	970	34.2	582	2 US-09-915-181A-4	Sequence 4, Appli
10	955	33.7	585	2 US-09-740-041-2	Sequence 2, Appli
11	941	33.2	567	2 US-09-949-016-11354	Sequence 11354, A
12	940	33.1	560	1 US-08-647-484-2	Sequence 2, Appli
13	940	33.1	560	1 US-08-647-484-2	Sequence 2, Appli
14	940	33.1	560	1 US-08-430-033A-2	Sequence 2, Appli
15	940	33.1	560	1 PCT-US96-05792-2	Sequence 2, Appli
16	936	33.0	560	1 US-08-805-118-4	Sequence 4, Appli
17	936	33.0	560	1 US-09-391-958-4	Sequence 4, Appli
18	936	33.0	560	1 US-09-915-181A-5	Sequence 5, Appli
19	935	33.0	850	2 US-09-915-181A-3	Sequence 3, Appli
20	876	30.9	497	2 US-09-949-016-6616	Sequence 6616, Ap
21	810	28.6	552	2 US-09-270-767-45540	Sequence 45540, A
22	808	28.5	563	2 US-09-915-181A-6	Sequence 6, Appli
23	808	28.5	576	2 US-08-864-785-1	Sequence 1, Appli
24	806.5	28.4	465	2 US-09-915-181A-8	Sequence 8, Appli
25	801.5	28.3	467	1 US-08-805-118-3	Sequence 3, Appli
26	801.5	28.3	467	2 US-09-391-958-3	Sequence 3, Appli

27	793	28.0	480	1 US-08-724-394A-9	Sequence 9, Appli
28	686	24.2	436	2 US-09-949-016-11448	Sequence 11448, A
29	686	24.2	470	1 US-08-724-394A-10	Sequence 10, Appli
30	615.5	21.7	401	1 US-08-805-118-1	Sequence 1, Appli
31	615.5	21.7	401	2 US-09-391-958-1	Sequence 1, Appli
32	592.5	20.9	480	1 US-08-724-394A-11	Sequence 11, Appli
33	562	19.8	380	2 US-09-949-016-7053	Sequence 7053, Ap
34	453	16.0	460	2 US-09-489-039A-9663	Sequence 9663, Ap
35	445.5	15.7	234	2 US-09-270-767-45527	Sequence 45527, A
36	391	13.8	436	2 US-10-154-419-98	Sequence 98, Appli
37	387	13.6	438	2 US-10-154-419-95	Sequence 95, Appli
38	375	13.2	186	2 US-09-270-767-45384	Sequence 45384, A
39	373	13.2	434	2 US-09-489-039A-13633	Sequence 13633, A
40	370	13.0	470	2 US-09-328-352-6912	Sequence 6912, Ap
41	369	13.0	439	2 US-09-172-952-14	Sequence 14, Appli
42	369	13.0	439	2 US-09-922-501-12	Sequence 12, Appli
43	368	13.0	157	2 US-09-270-767-61037	Sequence 61037, A
44	355.5	12.5	455	2 US-09-489-039A-9942	Sequence 9942, Ap
45	347	12.2	204	2 US-10-104-047-2253	Sequence 2253, Ap

ALIGNMENTS

RESULT 1
US-09-359-167-8
Sequence 8, Application US/09359167
Patent No. 6803448
GENERAL INFORMATION:
APPLICANT: Helleberg, Carl
APPLICANT: Fu, Changlin
TITLE OF INVENTION: GBS Toxin Receptor
FILE REFERENCE: CARB-008/01US
CURRENT APPLICATION NUMBER: US/09/359,167
EARLIER FILING DATE: 1999-07-21
EARLIER APPLICATION NUMBER: 60-693,843
NUMBER OF SEQ ID NOS: 12
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 8
LENGTH: 536
TYPE: PRT
ORGANISM: Homo sapiens
US-09-359-167-8

Query Match 100.0%; Score 2836; DB 2; Length 536;
Best Local Similarity 100.0%; Pred. No. 5.6e-295;
Matches 536; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MAGAMTPPPVQPARPGFGLSGRRSLCOVASTPAHVGVMSPVVDLARNDEESTDR	60
DB	1	MAGAMTPPPVQPARPGFGLSGRRSLCOVASTPAHVGVMSPVVDLARNDEESTDR	60
QY	61	TPLLPGAPRAEAAPVCCSARYNLAIAFGFFIVYALRVLSVALYDMVDNNTLLEENRT	120
DB	61	TPLLPGAPRAEAAPVCCSARYNLAIAFGFFIVYALRVLSVALYDMVDNNTLLEENRT	120
QY	121	SKACPEHSADIKYHNQTKKRYQMDAETQCMILGSPFYGIITIQIPGGYASKIGGMML	180
DB	121	SKACPEHSADIKYHNQTKKRYQMDAETQCMILGSPFYGIITIQIPGGYASKIGGMML	180
QY	181	GFGLIGRAVTLTTPRIADIGVPLTYRLBEGISGVTPPAHAMSSNAAPLERSKL	240
DB	181	GFGLIGRAVTLTTPRIADIGVPLTYRLBEGISGVTPPAHAMSSNAAPLERSKL	240
QY	241	SISVAGAGLQGVISLPLSGIICVYMMWTVYFFFGITIGFWMFLMTLWVSDTPQKHRS	300
DB	241	SISVAGAGLQGVISLPLSGIICVYMMWTVYFFFGITIGFWMFLMTLWVSDTPQKHRS	300
QY	301	HYEKEYLLSLRNQLSQKSVPVVPIKSLPLMAIVAAHSYMMFTTLLPTWKET	360
DB	301	HYEKEYLLSLRNQLSQKSVPVVPIKSLPLMAIVAAHSYMMFTTLLPTWKET	360

```
QY 361 LRFNVOENGFLSSLPYLGSWLCMLTSGOADNLRKKNFSTLCVRIFSLIGMIGPAVL 420
| | | | |
Db 361 LRFNVOENGFLSSLPYLGSWLCMLTSGOADNLRKKNFSTLCVRIFSLIGMIGPAVL 420
QY 421 VAAAGFICDYSLAVAFLLTISTTLGGFCSSGFSINHLDIAPSYAGILLGINTFATIGMV 480
| | | | |
Db 421 VAAAGFICDYSLAVAFLLTISTTLGGFCSSGFSINHLDIAPSYAGILLGINTFATIGMV 480
QY 481 GPVIAKSLTPDNTVGEQOTVFYIAAINVGAIFFTLPACGEVONMNLNDHGHRRH 536
| | | | |
Db 481 GPVIAKSLTPDNTVGEQOTVFYIAAINVGAIFFTLPACGEVONMNLNDHGHRRH 536
```

RESULT 2

```
US-09-949-016-7705
; Sequence 7705, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 7705
; LENGTH: 585
; TYPE: PR1
; ORGANISM: Human
US-09-949-016-7705
```

```
Query Match 98.9%; Score 2804.5; DB 2; Length 585;
Best Local Similarity 95.9%; Pred. No. 1.6e-291;
Matches 534; Conservative 1; Mismatches 1; Indels 21; Gaps 1;
```

```
QY 1 MAAGAMT-----PPRPVQPARPGGFGSLGRRSLTQVASTPAHV 39
| | | | |
Db 29 MAAGABARPPPLGGTAGTRRGRAVSSPPRPVQPARPGGFGSLGRRSLTQVASTPAHV 88
QY 40 GVMRSPPVRLAANDGESESTDRTPLLPGAPRAEAAVCCSARYNLAIIAFGFFIVYALRV 99
| | | | |
Db 89 GVMRSPPVRLAANDGESESTDRTPLLPGAPRAEAAVCCSARYNLAIIAFGFFIVYALRV 148
QY 100 NLSVALVDMVDSNTTLENDNTSKACPEHSAPIKVHHNOTGKKYQMDAETQGMILGSFFYG 159
| | | | |
Db 149 NLSVALVDMVDSNTTLENDNTSKACPEHSAPIKVHHNOTGKKYQMDAETQGMILGSFFYG 208
QY 160 YIIITDIPGVYASKIGKMLGFGILGTAVALTLFTPIAADLGVPPLVLRALBGLBGEVT 219
| | | | |
Db 209 YIIITDIPGVYASKIGKMLGFGILGTAVALTLFTPIAADLGVPPLVLRALBGLBGEVT 268
QY 220 PPAHAMSSWAPPLERSKLSISYAGAQGTIVISLPLSGIICYMMNTYVFFFGTIGI 279
| | | | |
Db 269 PPAHAMSSWAPPLERSKLSISYAGAQGTIVISLPLSGIICYMMNTYVFFFGTIGI 328
QY 280 FWFLLMTLVSDTPQKHRIKSHYEKEYILSLRNQSSOKSVPMVPLIKSLPLMAIVAAH 339
| | | | |
Db 329 FWFLLMTLVSDTPQKHRIKSHYEKEYILSLRNQSSOKSVPMVPLIKSLPLMAIVAAH 388
QY 340 FSYNMTFTYTLTLPTMYKEILRFNVOENGFLSSLPYLGSWLCMLTSGOADNLRKKNF 399
| | | | |
Db 389 FSYNMTFTYTLTLPTMYKEILRFNVOENGFLSSLPYLGSWLCMLTSGOADNLRKKNF 448
QY 400 STLCVRIRIFSLIGMIGPAVLVAAAGFICDYSLAVAFLLTISTTLGGFCSSGFSINHLDI 459
| | | | |
Db 449 STLCVRIRIFSLIGMIGPAVLVAAAGFICDYSLAVAFLLTISTTLGGFCSSGFSINHLDI 508
```

```
QY 460 PSYAGILLGINTTNPATIPGMVGVIAKSLTPDNTVGEQOTVFYIAAINVGAIFFTLFA 519
| | | | |
Db 509 PSYAGILLGINTTNPATIPGMVGVIAKSLTPDNTVGEQOTVFYIAAINVGAIFFTLFA 568
QY 520 KGEVQNMALNDHGHRRH 536
| | | | |
Db 569 KGEVQNMALNDHGHRRH 585
```

RESULT 3

```
US-09-359-167-2
; Sequence 2, Application US/09359167
; Patent No. 6803448
; GENERAL INFORMATION:
; APPLICANT: Hellerqvist, Carl
; APPLICANT: Fu, Changlin
; TITLE OF INVENTION: GBS Toxin Receptor
; FILE REFERENCE: CARB-008/01US
; CURRENT APPLICATION NUMBER: US/09/359,167
; CURRENT FILING DATE: 1999-07-21
; EARLIER APPLICATION NUMBER: 60-693,843
; EARLIER FILING DATE: 1998-07-22
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 495
; TYPE: PR1
; ORGANISM: Homo sapiens
US-09-359-167-2
```

```
Query Match 92.4%; Score 2621; DB 2; Length 495;
Best Local Similarity 100.0%; Pred. No. 5.8e-272;
Matches 495; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 42 MRSPPVRLAANDGESESTDRTPLLPGAPRAEAAVCCSARYNLAIIAFGFFIVYALRVNL 101
| | | | |
Db 1 MRSPPVRLAANDGESESTDRTPLLPGAPRAEAAVCCSARYNLAIIAFGFFIVYALRVNL 60
QY 102 SVALLVDMVDSNTTLENDNTSKACPEHSAPIKVHHNOTGKKYQMDAETQGMILGSFFYGI 161
| | | | |
Db 61 SVALLVDMVDSNTTLENDNTSKACPEHSAPIKVHHNOTGKKYQMDAETQGMILGSFFYGI 120
QY 162 ITQIPGVYASKIGKMLGFGILGTAVALTLFTPIAADLGVPPLVLRALBGLBGEVTP 221
| | | | |
Db 121 ITQIPGVYASKIGKMLGFGILGTAVALTLFTPIAADLGVPPLVLRALBGLBGEVTP 180
QY 222 AMEAMSSWAPPLERSKLSISYAGAQGTIVISLPLSGIICYMMNTYVFFFGTIGI 281
| | | | |
Db 181 AMEAMSSWAPPLERSKLSISYAGAQGTIVISLPLSGIICYMMNTYVFFFGTIGI 240
QY 282 FLIMTIVSDTPQKHRIKSHYEKEYILSLRNQSSOKSVPMVPLIKSLPLMAIVAAHFS 341
| | | | |
Db 241 FLIMTIVSDTPQKHRIKSHYEKEYILSLRNQSSOKSVPMVPLIKSLPLMAIVAAHFS 300
QY 342 YNMTFTYTLTLPTMYKEILRFNVOENGFLSSLPYLGSWLCMLTSGOADNLRKKNFST 401
| | | | |
Db 301 YNMTFTYTLTLPTMYKEILRFNVOENGFLSSLPYLGSWLCMLTSGOADNLRKKNFST 360
QY 402 LCVRIRIFSLIGMIGPAVLVAAAGFICDYSLAVAFLLTISTTLGGFCSSGFSINHLDIAPS 461
| | | | |
Db 361 LCVRIRIFSLIGMIGPAVLVAAAGFICDYSLAVAFLLTISTTLGGFCSSGFSINHLDIAPS 420
QY 462 YAGILGINTTNPATIPGMVGVIAKSLTPDNTVGEQOTVFYIAAINVGAIFFTLFAKG 521
| | | | |
Db 421 YAGILGINTTNPATIPGMVGVIAKSLTPDNTVGEQOTVFYIAAINVGAIFFTLFAKG 480
QY 522 EVQNMALNDHGHRRH 536
| | | | |
Db 481 EVQNMALNDHGHRRH 495
```

RESULT 4

```
US-09-915-181A-7
; Sequence 7, Application US/09915181A
; Patent No. 6818391
; GENERAL INFORMATION:
; APPLICANT: EDWARDS, ROBERT
; APPLICANT: BELLOCCHIO, ELIZABETH
; APPLICANT: FREMEAUX, ROBERT
; APPLICANT: REIMER, RICHARD
; TITLE OF INVENTION: NOVEL GLUTAMATE TRANSPORTERS
; FILE REFERENCE: 305T-932610US
; CURRENT APPLICATION NUMBER: US/09/915,181A
; CURRENT FILING DATE: 2002-03-26
; PRIOR APPLICATION NUMBER: US 60/220,556
; PRIOR FILING DATE: 2000-07-25
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 7
; LENGTH: 495
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-915-181A-7

Query Match          92.4%; Score 2621; DB 2; Length 495;
Best Local Similarity 100.0%; Pred. No. 5.8e-272;
Matches 495; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 MRSVPDLANDDEESTDRTPLLPGARAPAPCCSARYNLAALAFPGFPIYALAVNL 101
DB 1 MRSVPDLANDDEESTDRTPLLPGARAPAPCCSARYNLAALAFPGFPIYALAVNL 60
QY 102 SVALVDMVDSNTLTLEDNRSTKACPEHSAPIKVHHNQTGKKYQMDAEIQGNILGSFFYGYI 161
DB 61 SVALVDMVDSNTLTLEDNRSTKACPEHSAPIKVHHNQTGKKYQMDAEIQGNILGSFFYGYI 120
QY 162 ITQIPGGYVASKIGKMLGFGILGTAVALTLFTPIAADLGVPPLIVLRALGEGVTFP 221
DB 121 ITQIPGGYVASKIGKMLGFGILGTAVALTLFTPIAADLGVPPLIVLRALGEGVTFP 180
QY 222 AMHAMSSMAPPLERSKLSISYAGOLGTVISLPLSGIICYYNMWTVYVFFFTIGIFW 281
DB 181 AMHAMSSMAPPLERSKLSISYAGOLGTVISLPLSGIICYYNMWTVYVFFFTIGIFW 240
QY 282 FILMIMVSDTPQKHRIHYEKEYILSLRNQLSQKSPWVPIKSLPLMAIVAHFS 341
DB 241 FILMIMVSDTPQKHRIHYEKEYILSLRNQLSQKSPWVPIKSLPLMAIVAHFS 300
QY 342 YNMTFYTLTLPTMYKEILRFNVQENGFLSLPYLGSLCMILSGQAADNLRKXNFT 401
DB 301 YNMTFYTLTLPTMYKEILRFNVQENGFLSLPYLGSLCMILSGQAADNLRKXNFT 360
QY 402 LCVRRITSLIGMTGRAVFLVAGFIGDYSLAVALFISTITLGGFCSSGSINHLDIAPS 461
DB 361 LCVRRITSLIGMTGRAVFLVAGFIGDYSLAVALFISTITLGGFCSSGSINHLDIAPS 420
QY 462 YAGILGINTTFAITPGMGPVIAKSLTPNTVGEQTVFYIAAIVFQAIFFTLPAKG 521
DB 421 YAGILGINTTFAITPGMGPVIAKSLTPNTVGEQTVFYIAAIVFQAIFFTLPAKG 480
QY 522 EVQNMALNDHGHHRH 536
DB 481 EVQNMALNDHGHHRH 495

RESULT 5
US-09-359-167-4
; Sequence 4, Application US/09359167
; Patent No. 6803448
; GENERAL INFORMATION:
; APPLICANT: Hellerqvist, Carl
; APPLICANT: Fu, Changlin
; TITLE OF INVENTION: GBS Toxin Receptor
; FILE REFERENCE: CARB-008/01US
; CURRENT APPLICATION NUMBER: US/09/359,167
; EARLIER FILING DATE: 1998-07-22
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 10
; LENGTH: 495
; TYPE: PRT
; ORGANISM: Artificial Sequence
US-09-359-167-4
```

```
CURRENT FILING DATE: 1999-07-21
; EARLIER APPLICATION NUMBER: 60-693,843
; EARLIER FILING DATE: 1998-07-22
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 4
; LENGTH: 495
; TYPE: PRT
; ORGANISM: Ovis sp.
US-09-359-167-4

Query Match          82.1%; Score 2329; DB 2; Length 495;
Best Local Similarity 86.3%; Pred. No. 1.2e-240;
Matches 427; Conservative 39; Mismatches 29; Indels 0; Gaps 0;

QY 42 MRSVPDLANDDEESTDRTPLLPGARAPAPCCSARYNLAALAFPGFPIYALAVNL 101
DB 1 MRSVPDLANDDEESTDRTPLLPGARAPAPCCSARYNLAALAFPGFPIYALAVNL 60
QY 102 SVALVDMVDSNTLTLEDNRSTKACPEHSAPIKVHHNQTGKKYQMDAEIQGNILGSFFYGYI 161
DB 61 SVALVDMVDSNTLTLEDNRSTKACPEHSAPIKVHHNQTGKKYQMDAEIQGNILGSFFYGYI 120
QY 162 ITQIPGGYVASKIGKMLGFGILGTAVALTLFTPIAADLGVPPLIVLRALGEGVTFP 221
DB 121 ITQIPGGYVASKIGKMLGFGILGTAVALTLFTPIAADLGVPPLIVLRALGEGVTFP 180
QY 222 AMHAMSSMAPPLERSKLSISYAGOLGTVISLPLSGIICYYNMWTVYVFFFTIGIFW 281
DB 181 AMHAMSSMAPPLERSKLSISYAGOLGTVISLPLSGIICYYNMWTVYVFFFTIGIFW 240
QY 282 FILMIMVSDTPQKHRIHYEKEYILSLRNQLSQKSPWVPIKSLPLMAIVAHFS 341
DB 241 FILMIMVSDTPQKHRIHYEKEYILSLRNQLSQKSPWVPIKSLPLMAIVAHFS 300
QY 342 YNMTFYTLTLPTMYKEILRFNVQENGFLSLPYLGSLCMILSGQAADNLRKXNFT 401
DB 301 YNMTFYTLTLPTMYKEILRFNVQENGFLSLPYLGSLCMILSGQAADNLRKXNFT 360
QY 402 LCVRRITSLIGMTGRAVFLVAGFIGDYSLAVALFISTITLGGFCSSGSINHLDIAPS 461
DB 361 LCVRRITSLIGMTGRAVFLVAGFIGDYSLAVALFISTITLGGFCSSGSINHLDIAPS 420
QY 462 YAGILGINTTFAITPGMGPVIAKSLTPNTVGEQTVFYIAAIVFQAIFFTLPAKG 521
DB 421 YAGILGINTTFAITPGMGPVIAKSLTPNTVGEQTVFYIAAIVFQAIFFTLPAKG 480
QY 522 EVQNMALNDHGHHRH 536
DB 481 EVQNMALNDHGHHRH 495

RESULT 6
US-09-359-167-10
; Sequence 10, Application US/09359167
; Patent No. 6803448
; GENERAL INFORMATION:
; APPLICANT: Hellerqvist, Carl
; APPLICANT: Fu, Changlin
; TITLE OF INVENTION: GBS Toxin Receptor
; FILE REFERENCE: CARB-008/01US
; CURRENT APPLICATION NUMBER: US/09/359,167
; EARLIER FILING DATE: 1998-07-22
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 10
; LENGTH: 495
; TYPE: PRT
; ORGANISM: Artificial Sequence
US-09-359-167-10
```

```
Query Match      59.9%; Score 1698; DB 2; Length 495;
Best Local Similarity 67.8%; Pred. No. 6.2e-173;
Matches 335; Conservative 0; Mismatches 159; Indels 0; Gaps 0;

QY 42 MRSPPVDLANDEESTDRTPLPGAPRAEAPVCSARYNLA1LAEFGFIVYALRVNL 101
DB 1 MXXPVADAXXGEXXDRXXXXXXARXEXAPCCSARYNXXALXFFGFFXXYLKVLNL 60
QY 102 SVALVDMVDSNTTLEDRNRTSKACPEHSAPIKVHHNQTGKKYQMDAETQGMILGSPFYGYI 161
DB 61 XVXXVXMXDSXTTXDXNRKXSKCXEHSAPIKVVXXQTXGKXXWDATQGMILKXFFXXGYI 120
QY 162 ITQIPGGYVASKIGGKMLGFGILGTAVALTFTPIADLGVGPLIVLRALGEGGVTPP 221
DB 121 ITXIPGGYVASKXGGXXLXGXIXXXATLFTPPAADXGKXXXXLXALEXLEGXTXP 180
QY 222 AMHAMSSWAPPLERSKLSISYAGAQLTGTVISLPLSGIICYYMMNTYVYFFGTIGIFW 281
DB 181 AMHAMSSWAPPLERSKLSIXIXYAGAXLGTIVSLPLSGIICYYMMNTYVYFFXXGXXW 240
QY 282 FLIMVLVSDTPQKHRIKSHYEKEYILSSLRNOLSSQKVPWVPIKLSLPLMAIVAHFS 341
DB 241 FXXMIXLVSTPPXHXKXXXXXEXKXILSSLNQSSQKSVWXXXXKXLPMAXXVAXFS 300
QY 342 YNMTFTYTLTLLPTYMKEILRFNVQENGFLSLPYLGSWLCMILSGQAADNLRANKNFST 401
DB 301 YNMTFTYTLXLLPTXMKXXLRFXQENGFLSXKPYLXXMCMILKGAADNLRANKNFST 360
QY 402 LCVRIEFLSLIGMIGPAVFLVAGFIGCDYSLAVAFLLTISTLLGFCSSGFSINHLDIAPS 461
DB 361 XVXXRFXSLIXMIGPXXFLXXXXXXGCDYXKXVFLXISTXLGFCSSGFSINHLDIAPS 420
QY 462 YAGILGINTPTATPGMGVPIAKSLTDPNTVGEQVTVYIAAIVPFAIFFTLPFAGK 521
DB 421 YAGXLLGITNXFPATIXGMXGPXIXXSTPXNTXGEMQXFXXAAXNKFALFTLPFAGK 480
QY 522 EVONMALNDHGHGR 535
DB 481 EXONMXXDXDHXGR 494

RESULT 7
US-09-359-167-12
; Sequence 12; Application US/09359167
; Patent No. 6803448
; GENERAL INFORMATION:
; APPLICANT: Helicovera, Carl
; APPLICANT: Fu, Changlin
; TITLE OF INVENTION: GBS Toxin Receptor
; FILE REFERENCE: CARB-008/01US
; CURRENT APPLICATION NUMBER: US/09/359,167
; CURRENT FILING DATE: 1999-07-21
; EARLIER APPLICATION NUMBER: 60-693,843
; EARLIER FILING DATE: 1998-07-22
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 12
; LENGTH: 495
; TYPE: PRT
; ORGANISM: Artificial Sequence
US-09-359-167-12

Query Match      59.9%; Score 1698; DB 2; Length 495;
Best Local Similarity 67.8%; Pred. No. 6.2e-173;
Matches 335; Conservative 0; Mismatches 159; Indels 0; Gaps 0;

QY 42 MRSPPVDLANDEESTDRTPLPGAPRAEAPVCSARYNLA1LAEFGFIVYALRVNL 101
DB 1 MXXPVADAXXGEXXDRXXXXXXARXEXAPCCSARYNXXALXFFGFFXXYLKVLNL 60
QY 102 SVALVDMVDSNTTLEDRNRTSKACPEHSAPIKVHHNQTGKKYQMDAETQGMILGSPFYGYI 161
DB 61 XVXXVXMXDSXTTXDXNRKXSKCXEHSAPIKVVXXQTXGKXXWDATQGMILKXFFXXGYI 120
```

```
QY 162 ITQIPGGYVASKIGGKMLGFGILGTAVALTFTPIADLGVGPLIVLRALGEGGVTPP 221
DB 121 ITXIPGGYVASKXGGXXLXGXIXXXATLFTPPAADXGKXXXXLXALEXLEGXTXP 180
QY 222 AMHAMSSWAPPLERSKLSISYAGAQLTGTVISLPLSGIICYYMMNTYVYFFGTIGIFW 281
DB 181 AMHAMSSWAPPLERSKLSIXIXYAGAXLGTIVSLPLSGIICYYMMNTYVYFFXXGXXW 240
QY 282 FLIMVLVSDTPQKHRIKSHYEKEYILSSLRNOLSSQKVPWVPIKLSLPLMAIVAHFS 341
DB 241 FXXMIXLVSTPPXHXKXXXXXEXKXILSSLNQSSQKSVWXXXXKXLPMAXXVAXFS 300
QY 342 YNMTFTYTLTLLPTYMKEILRFNVQENGFLSLPYLGSWLCMILSGQAADNLRANKNFST 401
DB 301 YNMTFTYTLXLLPTXMKXXLRFXQENGFLSXKPYLXXMCMILKGAADNLRANKNFST 360
QY 402 LCVRIEFLSLIGMIGPAVFLVAGFIGCDYSLAVAFLLTISTLLGFCSSGFSINHLDIAPS 461
DB 361 XVXXRFXSLIXMIGPXXFLXXXXXXGCDYXKXVFLXISTXLGFCSSGFSINHLDIAPS 420
QY 462 YAGILGINTPTATPGMGVPIAKSLTDPNTVGEQVTVYIAAIVPFAIFFTLPFAGK 521
DB 421 YAGXLLGITNXFPATIXGMXGPXIXXSTPXNTXGEMQXFXXAAXNKFALFTLPFAGK 480
QY 522 EVONMALNDHGHGR 535
DB 481 EXONMXXDXDHXGR 494

RESULT 8
US-09-740-041-4
; Sequence 4; Application US/09740041
; Patent No. 6562593
; GENERAL INFORMATION:
; APPLICANT: MERKULOV, Karl et al
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL001001
; CURRENT APPLICATION NUMBER: US/09/740,041
; CURRENT FILING DATE: 2000-12-20
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 578
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-740-041-4

Query Match      34.2%; Score 970; DB 2; Length 578;
Best Local Similarity 40.9%; Pred. No. 1e-94;
Matches 201; Conservative 93; Mismatches 165; Indels 32; Gaps 9;

QY 51 RNDGEESTDRTPLPGAP---RAEAPVC-GSA-----RYNLALIAFGFFIVYALRVNL 101
DB 33 KQDNRETIETLE--DGKPLEVPEKAPLDCDCTCFGLPRRYIIAIMGSLGFCISFGICRNL 90
QY 102 SVALVDMVDSNTTLEDRNRTSKACPEHSAPIKVHHNQTGKKYQMDAETQGMILGSPFYGYI 161
DB 91 GVAIVDMVNNSTI---HRGKVIKEKA-----KFNWDPETVGMIHGSPFWGYI 135
QY 162 ITQIPGGYVASKIGGKMLGFGILGTAVALTFTPIADLGVGPLIVLRALGEGGVTPP 221
DB 136 ITQIPGGYVASKSLAANRVFGAAILTLSTLMLIPSAARHYGCVITVRILQGLVEGVTPP 195
QY 222 AMHAMSSWAPPLERSKLSISYAGAQLTGTVISLPLSGIICYYMMNTYVYFFGTIGIFW 281
DB 196 ACHGINSWAPPLERSRLATTSFCGSYAGAVIAMPLAGILVQYTGSSVFFYVYVGSFGMW 255
QY 282 FLIMVLVSDTPQKHRIKSHYEKEYILSSL---RNLSSQK---SVPWVPIKLSLPLMAIV 336
DB 256 YMFMLLVSESPAKHPTITDEERYIEESIGESANLLGAMEXKPTWPKRFTSPMPYALII 315
```

```

QY 337 VAFHSYNNMFYYTLTLTLPTYYKELAEFNNOENFSLPLATGMLCMILSGAADNIRAK 356
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 316 VANNFCRSMTFYLLILISQPAYFEEVGFGEISKVGLSAVPHLVNTIIVPIGGQADFIRSK 375
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY 357 WNFSTLCVRRIFSLIGMIPAVFLVAGFIGDCYSLAVALFTLSTLIGFCSSGFSINHL 456
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 376 QILSTTYRKAKNNCGFGEATILLVGY-SHTRGVALISFLVAVGSSGALISFANNHL 434
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY 457 DIASVAGTLGLTNTPATIPGVGVVIAKSLTPDNTVGEWQTFVYIAAINVFGAIFT 516
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 435 DIAPVYASIIAMGISNGVTLGWCPIIVGAMTKNKSREEMQVFLPALVHGVIFFYA 494
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
QY 517 LFAKEGVQWMA 527
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 495 LFAEGEKOPMA 505
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||

```

```

, RESULT 9
, US-09-915-181A-4
, Sequence 4, Application US/09915181A
, Patent No. 6818391
, GENERAL INFORMATION:
, APPLICANT: EDWARDS, ROBERT
, APPLICANT: BELLOCCHIO, ELIZABETH
, APPLICANT: FREMEAUX, ROBERT
, APPLICANT: REIMER, RICHARD
, TITLE OF INVENTION: NOVEL GLUTAMATE TRANSPORTERS
, FILE REFERENCE: 305T-932610US
, CURRENT APPLICATION NUMBER: US/09/915,181A
, CURRENT FILING DATE: 2002-03-26
, PRIOR APPLICATION NUMBER: US 60/220,556
, PRIOR FILING DATE: 2000-07-25
, NUMBER OF SEQ ID NOS: 11
, SOFTWARE: PatentIn version 3.0
, SEQ ID NO 4
, LENGTH: 582
, TYPE: PRT
, ORGANISM: Rattus rattus
, US-09-915-181A-4

```

Query Match	34.2%	Score 970;	DB 2;	Length 582;
Best Local Similarity	40.9%;	Pred. No. 1e-94;		
Matches 201; Conservative	93;	Mismatches 165;	Indels 32;	Gaps 9

QY	51	ANDGEESIDRPLRLGAP---	RABAPVC-GSA-----	RYNLAIALAFPGEPITYALRYNL	101
Db	34	KODNRREITELTE--	DGKPLEVPEKCAPLDDCTCGFLPRRYIILAINSGLPJCSIFGRCNL		91
QY	102	SVALDAMDVDSNTLTEDNRITSKACPEHSAP	IKHHNOTGKKYQOMDAETQOMILGSPFYGYI	161	
Db	92	GVALIDVWNNSTI--	HGKGKVIKKA-----	KFNMDPEITVGMHGSFFWGYI	136
QY	162	ITQIPGVYVASKIGGKMLGFGILGTAVILTF	TPPIADIGVGPLVLRALEGLGEGVTEP	221	
Db	137	ITQIPGVYASRLAANRVFGAAILLTSTLNM	LIPSAARVHGCVIPFVRIIGLVEGVTEP	196	
QY	222	AMHAMSSWAPLERSEKSLISTISYAAGQ	GTVISLPSLGITCYMMNTYTFYFPGTIGITW	281	
Db	197	ACHGIMSWAPPLERSRLATTSFCGSA	YAGAVIAMPPLAGILVYTGSSVFYAYGSPGMW	256	
QY	282	FLWMIMLSDTPQKHKRI	SHYEKEYISL---	RNOLSQK--SVPMVFLIKSLPLPMATV	336
Db	257	YMFMLIVSESPAGHPIT	ITDEBRRIIEBISGSANLLGMERKTPMKRCEFTSMRYAII	316	
QY	337	VAHFSYNNTFYTLTLTLP	TYMKEILRFVQENGFLSLPEYLGSMCLISGQAADNLRKX	396	
Db	317	VANFCRSWTFYLLIS	OPAYFEVPGFEISKVGMLSAVBHLVMTIIVP	IGQIADPLRSK	376
QY	397	WNFSTLCYRRI	FLSLGIMGPAFLVLAAGTICDDYSLA	AFIITSTLGGFCSGSINHL	456
Db	377	QILSTTVYTRKIMNCG	FGMEATLLLVG--SHTRGVAISFLVLA	VGSGFALISGFVNHL	435

QY	457	DIAPFYAGLLGITYNTPATIPGMGVVIXKSLTPDNTVGEMQTFYIAAINVFGAFF	516
		: : : : : : : : :	
Db	436	DIAPFYASLIMQISNGVGTLSGMCIIIVGANTKKNSREMQVFLIALVHGVIFYA	495
		: : : : : : : : :	
QY	517	LFAPGEVQNWMA	527
		: : : : : : : : :	
Db	496	LFAPGEKQPMWA	506

```

RESULT 10
US-09-740-041-2
; Sequence 2, Application US/09740041
; Patent No. 6562593
; GENERAL INFORMATION:
; APPLICANT: MERRILOV, Karl et al
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: C1:001001
; CURRENT APPLICATION NUMBER: US/09/740,041
; CURRENT FILING DATE: 2000-12-20
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 589
; TYPE: PRT
; ORGANISM: Human
US-09-740-041-2

```

Query Match	33.7%;	Score 955;	DB 2;	Length 589;
Best Local Similarity	41.9%;	Pred. No. 4.3e-93;		
Matches 192; Conservative	88;	Mismatches 156;	Indels 22;	Gaps 7

Oy	76	CCS--ARNALMLAFEGFFIVALEVNLVSLVDVDSVTLTIEDNRTSKACEHSPAIKV	133
Db	69	CCGHPKRIIIMSLGFCISRGICRNLGVAIVENVNNSVYVDGK-----PE-----	116
Oy	134	HHNQTGGKYQWDAETQGMILGSPFYGYITITPGGVASKIGCKMLLFGILGTAVLTLP	193
Db	117	--IGTA--QPNMDPELVGLIHGSPFWGYIMTQIPGCFISNKPANRVFGAILELTSLNMF	173
Oy	194	TPIAADGVGLIYLRALEGGEQGTPTPAMAMMSMAPLERSKULSTSYAQLGCTVI	253
Db	174	IPSAARVHYGVCMCYRILIQLEGGVITYPACHGMSKMAPPLERSKRLATTSFGCSYAGAVV	233
Oy	254	SLPLSGIICYYMNTTYVYFEGTIGIFWFLIMLWVSDTPQKHRSISHEKEYIILSL--	311
Db	234	AMPLAGVLVOYIGMSSVFIYIGMGEIIMYMFLLDAYECPANPHITISNEKITYIETSIQ	293
Oy	312	-RNQLSSQK-SVPMVPIILKSILPLMAIVVAHFSYNMTFTLLTLPTMYKEILRPNVOENG	369
Db	294	GANVVASLTSKFTSPMKRPFSTSLPYVAIIIVANFCRSMTFYLLILSQAPAEFEVGFALSKVG	353
Oy	370	FLSLPLVGSWLMILSGQADNLRKANNFTLCVRILFSLIGMIGPAVFLVAAGICGD	429
Db	354	LLSAVPHVMWMTIVPIQGLADYLRKSROIITLTARKIMNCGFGMEATLLLVGG-SHT	412
Oy	430	YSLAVAFLLITSTLGGFCSSGSFINHLDIAPSYAGILGITTPTATIPGMVSPVIAKSLT	489
Db	413	KGVASIFLVLAVGSGFALISGFNNHDIARVYASILMGISNGVOTLSGMVPLIVGAMT	472
Oy	490	PDNTVGEHQVIFYIAALINVFGAIFTLFANGEVONMA	527
Db	473	RHKTRSEWONVFLIALVHSGVIFYGVAFASGEKQEWMA	510

RESULT 11
US-09-949-016-11354
; Sequence 11354; Application US/09949016
; Patent No. 681239
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

Db 253 LTVSYSPALHPSEISBEERKYEIDAIGESAKLMNPL-TKSTPMRRFTSMFYALIVAN 311
Qy 340 PSYNTFFYLLTLPTMYKEILRFNVOENGLSSLPYLSWLCMILSGAADNLRAKMF 399
Db 312 FCRSWTFYLLLIQPDYFEVFGFEISKVGLVSALPHLWMTIIVPIGGQIADFLRSRIM 371
Qy 400 STLCVRRIPLSLIGMIGPAVFLVAAFGICDYSILAVALFTISTLLGFCSSGFSINHLDA 459
Db 372 STTNVRKLMNCGFGMEATLLVVGY-SHSGVAISFLVAVGSGFAISGFVNHLDA 430
Qy 460 PSYAGILGTTNTFATIPGNVGVIAKSLTPDNTVGEMQTVFYIAAINVFGAIFFTLFA 519
Db 431 PRYASILMGISNGVGLSGVNCPIVGAAMTKHKTREEMQYVFLIASLVHVGVIYGVFA 490
Qy 520 KGEVONWA 527
Db 491 SGEKQPMWA 498
RESULT 15
PCT-US96-05792-2
; Sequence 2, Application PC/TUS9605792
; GENERAL INFORMATION:
; APPLICANT: Paul, Binhui
; TITLE OF INVENTION: HUMAN BRAIN SODIUM DEPENDENT INORGANIC
; TITLE OF INVENTION: PHOSPHATE COTRANSPORTER
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: United States of America
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/05792
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/430,033
; FILING DATE: April 27, 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Gaylo, Paul J.
; REGISTRATION NUMBER: 36,808
; REFERENCE/DOCKET NUMBER: X-10006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (317) 276-0756
; TELEFAX: (317) 276-3861
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 560 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US96-05792-2
Query Match 33.1%; Score 940; DB 5; Length 560;
Best Local Similarity 39.5%; Pred. No. 1.6e-91;
Matches 193; Conservative 93; Mismatches 162; Indels 40; Gaps 8;
Qy 53 DGEESTDRT---PLLP-----GAPRAEAPVCCSARYNLAIAPFGFFIVYALRVNLISVAL 105
Db 38 DGRPVTTQTDRDPVVDCTCFGLR-----RYIAIAMSGLGFCISFGIRCNLGVAI 87
Qy 106 VDWVDSNTTLEDNRTSKACPEHSAPRIKYNHNQTSKYQWDAETQGMILGSEFFYIITQI 165
Db 88 VSNVNNSTT-----HRGHHVVVQKQKQ-----FSWDEPVTGLIHGSEFFMGYIVTQI 132

Qy 166 PGGVYASKIGCKMILGFGILGTAVLTFTPIAADLGVPILVLALEGGEVTPPAMHA 225
Db 133 PGGFIOCKRAANRVFPFAIVASTLMMLPSAARVHYGCVIPRIIQLGVEGTYRACHG 192
Qy 226 MMSMAAPPLERSKTLISYAGAQLTVISLPLSGIICYYNMNTVYFFPGTIGIFWPLFM 285
Db 193 IWSKMAPPLERSRLATTAFCGSYAGAVVAMPPLAGVLVQSGMSVFYVGSFGIFWYLFM 252
Qy 286 IWLVDTPQKHRIISHYEKEYT-----LSSLRNQSSQSVWPVPIILKSLPMAIVAH 339
Db 253 LTVSYSPALHPSEISBEERKYEIDAIGESAKLMNPL-TKSTPMRRFTSMFYALIVAN 311
Qy 340 PSYNTFFYLLTLPTMYKEILRFNVOENGLSSLPYLSWLCMILSGAADNLRAKMF 399
Db 312 FCRSWTFYLLLIQPDYFEVFGFEISKVGLVSALPHLWMTIIVPIGGQIADFLRSRIM 371
Qy 400 STLCVRRIPLSLIGMIGPAVFLVAAFGICDYSILAVALFTISTLLGFCSSGFSINHLDA 459
Db 372 STTNVRKLMNCGFGMEATLLVVGY-SHSGVAISFLVAVGSGFAISGFVNHLDA 430
Qy 460 PSYAGILGTTNTFATIPGNVGVIAKSLTPDNTVGEMQTVFYIAAINVFGAIFFTLFA 519
Db 431 PRYASILMGISNGVGLSGVNCPIVGAAMTKHKTREEMQYVFLIASLVHVGVIYGVFA 490
Qy 520 KGEVONWA 527
Db 491 SGEKQPMWA 498

Search completed: June 7, 2006, 05:44:37
Job time : 53 secs

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2006, 05:38:01 ; Search time 42 Seconds
(without alignments)
1227.909 Million cell updates/sec

Title: US-10-823-506-8

Perfect score: 2836

Sequence: 1 MAGAMTPPPVQPARPGF.....LPAKGEVQNMALNDHGHHRH 536

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR_80:*
2: p1r1:*
3: p1r2:*
4: p1r3:*
5: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	987	34.8	493	2	G88553
2	967	34.1	472	2	S28286
3	936	33.0	560	2	I59302
4	822.5	29.0	465	2	A56410
5	809.5	28.5	465	2	S69915
6	808	28.5	563	2	T43650
7	808	28.5	576	2	H88548
8	801.5	28.3	467	2	A48916
9	797.5	28.1	573	2	T23589
10	786.5	27.7	465	2	I39473
11	767.5	27.1	544	2	T24633
12	747	26.3	512	2	H84658
13	629.5	22.2	501	2	B89135
14	562	19.8	413	2	T01534
15	531	18.7	537	2	T45634
16	530.5	18.7	466	2	S40767
17	492.5	17.4	561	2	D84800
18	486.5	17.2	592	2	S44742
19	476.5	16.8	562	2	T25419
20	460.5	16.2	516	2	T24729
21	452	15.9	530	2	T29418
22	445	15.7	380	2	T29248
23	444.5	15.3	543	2	T32496
24	432.5	15.2	499	2	T15201
25	431	15.2	506	2	T29968
26	419.5	14.8	493	2	T25357
27	392.5	13.8	478	2	T33942
28	389.5	13.7	445	2	T23590
29	388.5	13.7	462	2	T34365

30	386.5	13.6	473	2	T31717	hypothetical prote
31	380.5	13.4	425	2	A90055	hypothetical prote
32	372	13.1	420	2	S44900	ZK652.10 protein -
33	370.5	13.1	428	2	AD0675	membrane transport
34	362	12.8	485	2	T24115	hypothetical prote
35	360	12.7	455	1	H69752	probable glucarat
36	357.5	12.6	455	2	T34366	hypothetical prote
37	357	12.6	452	2	AE0861	probable glucarate
38	355.5	12.5	659	2	T33557	hypothetical prote
39	353	12.4	516	2	T27092	hypothetical prote
40	349	12.3	450	2	A65061	probable glucarate
41	349	12.3	450	2	C85930	probable transport
42	349	12.3	450	2	A91085	probable transport
43	339.5	12.0	445	1	D65171	hypothetical 48.8
44	336.5	11.9	493	2	T19383	hypothetical prote
45	334	11.8	487	2	T23776	hypothetical prote

ALIGNMENTS

RESULT 1

G88553 protein C38C10.2 [imported] - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C/Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004

C/Accession: G88553

R/Anonymous, The C. elegans Sequencing Consortium.

Science 282, 2012-2018, 1998

A/Title: Genome sequence of the nematode C. elegans: a platform for investigating biolo

A/Reference number: A75000; MUID:99069613; PMID:9851916

A/Note: see websites genome.wustl.edu/genec/C_elegans/ and www.sanger.ac.uk/Projects/C_el

A/Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; an

A/Accession: G88553

A/Status: Preliminary

A/Molecule type: DNA

A/Residues: 1-493 <STO>

A/Cross-references: UNIPROT:Q03567; UNIPARC:UP1000013B804; GB:chr_III; PIDN:CAA79549.1;

C/Genetic8:

A/Gene: C38C10.2

A/Map position: 3

Query Match	34.8%	Score 987	DB 2	Length 493
Best Local Similarity	39.7%	Pred. No. 2.6e-70		
Matches 194	Conservative 94	Mismatches 179	Indels 22	Gaps 6
QY	55	EESTDRTPLLPGAPRAAPVCCSARYNLAIAFFGFIVYALRVNLVALVDMVDSNTT	114	
DB	2	EGATTKEPLVP-----STRPALSLVNFPGCLVYMMRTNMSFAVVCVNNKNT	49	
QY	115	LEDNRISKACPEHSAPIKVHHNOTGKKYQWDATOGHILGSPFYGIITQIPGIVASKI	174	
DB	50	DTEGEKVSRCQKEMTPVENS SVIG--BFMDDKOTGVNLSPFYGYGSOIIGHILASRY	108	
QY	175	GKKMLIGFGILGTAVALTLFPPIAADGVGPLVIRALEGEGEVTFPAMAMMSSNAPL	234	
DB	109	GGRKVVVTVTLGSLTLPLNLPVPAKRSEYLAIRALIGLQGTATPAAHTMMSWGCPPL	168	
QY	235	ERSKLISISYAGQGLTVISLPLSGIICYY---MMNVYVFPGTIGIFWFLMIWLVD	291	
DB	169	ELSLVLTGVTVAAQIGNVILPLSGFLCEVGFQDGMGMSIFYLIGVFCVLTAAVVMVSSD	228	
QY	292	TPQKHRIKSHYKEKYLIISSLRNQLSSQ---KSPVNPILIKSLPMAIVVAHSYNTFY	347	
DB	229	KPAHPRIPTEEKQYIVTAVEASMGKDTGVPTSPWKILITSPAVMCMAGHAGDGAY	288	
QY	348	TLTLPTPYMKELIRFVQNGEFLSLPYLGSWLCMILSGQAADNLRAKNPFSTLCVARI	407	
DB	289	TMLVSLPSFLKDVLTGLNLSLGAVASIPYLAFLAINAGVADTLRSKQIISTLNTERRA	348	
QY	408	FSLIIGMIGPAVLVAAGFIGCDYS-LAAVLTISTTLGFCSSGGSFINHLDIAPSVAGIL	466	
DB	349	AMVALIIGGQIFLVASGYCCGQDVILVIFITCGMAISIGQYAGFVNVNLEIAPFSGTV	408	

Qy 134 HNNQTKKYOMDAETQGMILGSFFYGYIITQIPGGYVASKIGKMLGFGILGTAVALTLF 193
 Db 64 -DNTKPNVYVMSPPVOGIIIFSSIFYGAFLIOIPVIGYISGIKIKIGFALFLSSVISIF 122
 Qy 194 TPIAADLGVGLYLALBEGVTPPAMHAMSSNAPPLERSKLLISYAGOLGTVI 253
 Db 123 IPOAAVGEFTWIIICRVVOGITOCTVTAQHEIIVWKNAPPLERGLTSMISLGGPFI 182
 Qy 254 SLPLSGIICYYMMNTVYFPGTIGIFWFLIMWLVSBDTPQKHRIKSHYEKEYITLSSLRN 313
 Db 183 VLLVLTGICESTLGNPMVFIYFGACCAVCLLWFLYDDPDHPCVSLHEKEYITSSLIQ 242
 Qy 314 QLSG-OKSVPMVPIKSLPLMAIVAHFSYNMFTYTLTLPTMYKEILRFNVQENGLS 372
 Db 243 QGSSTROSLPIKMIKSLPLMAISFCGFAYIMTSLRLIVTPTLINSMLHNDIRENGLS 302
 Qy 373 SLPLSGWLMILSGQAADNLRAKNFSTLCVRRIISLIGMIGAVFLVAAAGTICDYSL 432
 Db 303 SLPLFAMICGVLAGHTADEFLMSRNMLSLTAIRKLTFAIGLLPIVFSMCLYLSSGFYS 362
 Qy 433 AVAFITSTLTGGFCSSGFSINHLDIAPSYAGILGTTNTPATIPGMVGPVIAKSLTPDN 492
 Db 363 TITFLIINASSSFCGLGALINADLAPRYVVFYKGVTTLIGMTGMTSTVAGLPLISQD 422
 Qy 493 TVGEMQTVFYIAAIVNFGAIFFTLFAKGEVONMALNDHH 532
 Db 423 PESSWFKIFLMSIINVISIFYLIFAKELIQDAKEKOH 462

RESULT 5

S69915
 sodium-phosphate transport system 1 - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 14-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
 C:Accession: S69915
 R:Chong, S.S.; Kozak, C.A.; Liu, L.; Kristjansson, K.; Dunn, S.T.; Bourdeau, J.E.; Hughes
 Am. J. Physiol. 268, 1038-1045, 1995
 A:Title: Cloning, genetic mapping, and expression analysis of a mouse renal sodium-deper
 A:Reference number: S69915
 A:Accession: S69915
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-465 <CHO>
 A:Cross-references: UNIPROT:Q61983; UNIPARC:UP1000002930D; EMBL:X77241; NID:G887522; PIR

Query Match 28.5%; Score 809.5; DB 2; Length 465;
 Best Local Similarity 36.7%; Pred. No. 2,7e-56;
 Matches 167; Conservative 90; Mismatches 189; Indels 9; Gaps 2;
 Qy 74 PVCCSARYNLAIAFGFFIVYALRVNLSVALVDVMSNTTLEDNRTSKACPEHSAPIK 133
 Db 11 PGCSFRYGAIIILHFNINIMAOVCNLITMAV-----NNTGSHLSNESVEM 62
 Qy 134 HNNQTKKYOMDAETQGMILGSFFYGYIITQIPGGYVASKIGKMLGFGILGTAVALTLF 193
 Db 63 LDVKKPVPVMSPPDIQGLILSVFPGVWVQAPVGYLSGIYPMKRITIGSSLFLSSLSL 122
 Qy 194 TPIAADLGVGLYLALBEGVTPPAMHAMSSNAPPLERSKLLISYAGOLGTVI 253
 Db 123 IPFAAVGALVIVCRVLOGIAQGTSTGQHEIIVWKNAPPLERGLTSMISLGGVMPFI 182
 Qy 254 SLPLSGIICYYMMNTVYFPGTIGIFWFLIMWLVSBDTPQKHRIKSHYEKEYITLSSLRN 313
 Db 183 VLLVLTGICDLGMVNFYIFGIYGCVLSLSWFLFLFDDPKDHPYSSSKDIIISLMQ 242
 Qy 314 QLSG-OKSVPMVPIKSLPLMAIVAHFSYNMFTYTLTLPTMYKEILRFNVQENGLS 372
 Db 243 QASSGROSPIKAMLSLPLMAIILNSFAFIMSNLSLVTPTPIFISVLHVNRENGLS 302
 Qy 373 SLPLSGWLMILSGQAADNLRAKNFSTLCVRRIISLIGMIGAVFLVAAAGTICDYSL 432
 Db 303 SLPLVLAIVICGILAGQSDFFLTRKIFSIYTVKRLFTTIGSCFVIFIMCLVLSNFFS 362

Qy 433 AVAFITSTLTGGFCSSGFSINHLDIAPSYAGILGTTNTPATIPGMVGPVIAKSLTPDN 492
 Db 363 TITFLIINASTLSYSGOLINADLAPRYVGPVIAKAVTALIMGGLISSTLADLIINQD 422
 Qy 493 TVGEMQTVFYIAAIVNFGAIFFTLFAKGEVONMA 527
 Db 423 PEVAMHKISPLMAGINTCLVFYFLFAKGEIQDMA 457

RESULT 6

T43650
 probable sodium-dependent inorganic phosphate cotransporter - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 09-Jul-2004
 C:Accession: T43650
 R:Lee, K.Y.N.; Sawin, E.R.; Chalfie, M.; Horvitz, H.R.; Avery, L.
 submitted to the EMBL Data Library, September 1998
 A:Description: EAT-4, a homolog of a mammalian sodium-dependent inorganic phosphate cot
 A:Reference number: 222599
 A:Accession: T43650
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-563 <LEE>
 A:Cross-references: UNIPROT:Q9T2N7; UNIPARC:UP1000007DE54; EMBL:AF095787; PIDN:AA64972
 A:Experimental source: strain Bristol N2
 C:Genetics:
 A:Gene: eat-4
 A:Map position: 3
 A:Function:
 A:Description: is necessary for glutamatergic neurotransmission

Query Match 28.5%; Score 808; DB 2; Length 563;
 Best Local Similarity 36.4%; Pred. No. 4.4e-56;
 Matches 166; Conservative 88; Mismatches 176; Indels 26; Gaps 7;
 Qy 77 CSARYNLAIAFGFFIVYALRVNLSVALVDVMSNTTLEDNRTSKACPEHSAPI-KVN 135
 Db 53 CRKRMILAILANMGFMISFCIRCNFGAKTHMYKNYTD-----PYGVHM 97
 Qy 136 NQTKKYOMDAETQGMILGSFFYGYIITQIPGGYVASKIGKMLGFGILGTAVALTLF 195
 Db 98 HE----FNMTIDELSVNESSYFYGYLTQIPAGFLAKFPNNKLPFGIGVGAFLNTLIP 153
 Qy 196 IAADLGVGLYLALBEGVTPPAMHAMSSNAPPLERSKLLISYAGOLGTVIS 254
 Db 154 YGKRVSDIYVATIQTQGLVQGVCPAMGVNRVMAPEWRSKLATTAFTGSTAGAVLG 213
 Qy 255 LPLSGIICYYMMNTVYFPGTIGIFWFLIMWLVSBDTPQKHRIKSHYEKEYI---LSSL 311
 Db 214 LPLSAFLVSVSMAAPVYLGVCVGMALIMFCVTEPKPAFHPTISQEEKIFIEDAIGHV 273
 Qy 312 RNOLSOQSVPMVPIKSLPLMAIVAHFSYNMFTYTLTLPTMYKEILRFNVQENGLF 371
 Db 274 SNHTPIRTSIPWKAIVTSKPVMAIIVANFARSWTFYLLQNLTYMKEALGMKIAOSGL 333
 Qy 372 SSLPYGSMWLMILSGQAADNLRAKNFSTLCVRRIISLIGMIGAVFLVAAAGTICDYS 431
 Db 334 AAIPLHMGCVVLMGQDLADYLSNKILSTTAVRKIFNCGGEGGAAPMLIYVYTTSD-T 382
 Qy 432 LAVAFITSTLTGGFCSSGFSINHLDIAPSYAGILGTTNTPATIPGMVGPVIAKSLTPD 491
 Db 393 TALMALIAVGMGFMISGNVNHLDIAPRYAAILMGFSIGITLGLTCPTTAEFTAH 452
 Qy 492 NTVGEWQTVFYIAAIVNFGAIFFTLFAKGEVONMA 527
 Db 453 SKHG-WTSVFLASLIHFTGVTFYAVYASGELOEMA 487

RESULT 7

H88548
 protein ZK512.6 (imported) - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans

Db 47 RMQIALAHFGFAISFGIRSNFGVAKRMNNFT-----DAYGEVHE---- 88

Qy 140 KKYQWDAETQGMILGSFFGYIITQIPGGYVASKIGKMLGPGILGTAVLTLPPIAD 199

Db 89 KEFFWTGETVGMWESSFFGYAASQIPAGVIAAFAPNKLPMILGILFASLINITATC- 146

Qy 200 LGVGP-----LIVRALEGEGVTPPAMHAMSSMAPPLERSKLLSISYAGOLGTVIS 254

Db 147 LNFHFPDIFPMVITQWQGLAGLGVCPAMHGVKMYMAPPLERSKLATTTFTGASGVWVG 206

Qy 255 LPLSGIICYYNMWTVYFFPGTIGIFWFLMIWLVSDTPQKHRIISHYEKEYI--LSSL 311

Db 207 LPASAVYLSHFMSSTPFFVFGALGIWISILWIVSGSPETHGYISADEKKYITREKGSV 266

Qy 312 RNOLSSQKSVPMVPIKSLPLMAIVAHFSYNMFYLLTLTPRYMEILRFNVOENGL 371

Db 267 AVKMTLTTLPRMWTSTAVMAIILCSFCRMSFFLLGNQLYMDVHLIDIKNSGLI 326

Qy 372 SSPLYSGLWCMILSGQAADNLRAKNFSTLCVRRISLIGMIGPAFLVAAGFIC-- 428

Db 327 AIRPQGMCIYTLTISQSLDYLRSCKMSTPAVKSVNTRG-----FTVEAVMLGCLAF 380

Qy 429 --DYSLAVALFTISTTLGFCSSGFSINHLDIAPSVAGILGINTPATIPGMVPIAK 486

Db 381 VRBPVIAVTEFLIACSGAGAVLSGFNVNHFIDIAFRHAPIIMGINGCAIAG-VGGIVTN 439

Qy 487 SLTPDNTVGEWQTVFYIAAIVNFGAIFFTLFAKEGVQNA 527

Db 440 SLTYQNPDG-WQWVFLMAISIDIFGILFPLFAKGDLVPA 479

RESULT 10

139473
Na+-dependent phosphate cotransporter - human
C/Species: Homo sapiens (man)
C/Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 09-Jul-2004
C/Accession: I39473
R: Miyamoto, K.; Tatemmi, S.; Sonoda, T.; Yamamoto, H.; Minami, H.; Taketani, Y.; Takeda, Biochem. J. 305, 81-85, 1995
A/Title: Cloning and functional expression of a Na(+)-dependent phosphate co-transporter
A/Reference number: I39473; MUID:95126933; PMID:7826357
A/Accession: I39473
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-465 <RES>
A/Cross-references: UNIPROT:Q14916; UNIPARC:UP1000016A56A; GB:ID28532; NID:9639841; PIDN:

Query Match

Best Local Similarity 37.7%; Score 786.5; DB 2; Length 465;
Pred. No. 1.8e-54;

Matches 177; Conservative 79; Mismatches 186; Indels 27; Gaps 6;

Qy 74 PUCSARYNLAIAAFGFFIVALARVNLVALVDVDSNTT--LEDNRTSKACEHSAPI 131

Db 11 PGCSFRYGLSFLVHCNVIITQIMCLNLTVMVNVNSTDPHGIPNTSTYKLLDNINPM 70

Qy 132 KYAHNQTQKKYQDAETQGMILGSFFGYIITQIPGGYVASKIGKMLGPGILGTAVLT 191

Db 71 -----YMSPDVQGIILSTSTYGVIIIOVPGVFSGIYSTKMGFALCLSSVLS 120

Qy 192 LFTPIADLGVPLIVLRALEGEGVTPPAMHAMSSMAPPLERSKLLSISYAGOLGT 251

Db 121 LILPRAAGIVAVVGVRAVQGAAGIVATAQFEIYKMAPPLERGLTSMSTISGFLGP 180

Qy 252 VISLPSGIICYYNMWTVYFFPGTIGIFWFLMIWLVSDTPQKHRIISHYEKEYIISL 311

Db 181 FYLVLTAVTGVICSLGPMVFIYIFGACCAVCLMFVLYDDPKDHPICISGEKEYITSSL 240

Qy 312 RNOL-SSQKSVPMVPIKSLPLMAIVAHFSYNMTFTLLTLPTYKKEILRFVQENGF 370

Db 241 VQGVSSRSQSLPIKAILIKSPWAISIGSFTFFSHIMTLVPMFINSMLHVNIKENGF 300

Qy 371 LSSLPIYGLSWLCLMSGOADNLRAKVNFSTLCVRRISLIGMIGPAFLVAAGFICDY 430

Db 301 LSSLPIYFAMICGNLAQSLDFLTRNLSVIAVRKLFTAAGFLPAIFGVLPLYSTP 360

Qy 431 SLAVALFTISTTLGFCSSGFSINHLDIAPSVAGILGINTPATIPGMVPIAKSLT- 489

Db 361 YSIVIFLILGANGSFCLGVPFLGIDIAPIRYFEFT---KACSTLGMIGGLIASTLTG 416

Qy 490 -----PDNTVGEWQTVFYIAAIVNFGAIFFTLFAKEGVQNA 532

Db 417 LILKQDESA---WFKTFILMAINVTGLFYLVATRAELQDAKEKQH 462

RESULT 11

124633
hypotheical protein T07A5.3 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C/Accession: T24633
R: Buck, D.
A/Submitted to the EMBL Data Library, February 1995
A/Reference number: Z19915
A/Accession: T24633
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-544 <WIL>
A/Cross-references: UNIPROT:Q10046; UNIPARC:UP1000013C01E; EMBL:Z48055; PIDN:CAA88134.1
A/Experimental source: clone T07A5
A/Genetics:
A/Gene: CESP:T07A5.3
A/Map position: 3
A/Intons: 38/3; 87/3; 272/3; 322/3; 444/3; 476/3; 496/1; 517/2

Query Match

Best Local Similarity 35.9%; Score 767.5; DB 2; Length 544;
Pred. No. 6.8e-53;

Matches 166; Conservative 89; Mismatches 164; Indels 43; Gaps 9;

Qy 80 RYNLAIAAFGFFIVALARVNLVALVDVDSNTTLEDNRTSKACEHSAPIKYAHNQTG 139

Db 47 RMQIALAHFGFAISFGIRSNFGVAKRMNNFT-----DAYGEVHE---- 88

Qy 140 KKYQWDAETQGMILGSFFGYIITQIPGGYVASKIGKMLGPGILGTAVLTLPPIA-- 197

Db 89 REFLMTGAEVGMWESSFFGYAASQIPAGVIAAFAPNKLPMILGILFASLINITATC- 146

Qy 198 ----ADLGVPLIVLRALEGEGVTPPAMHAMSSMAPPLERSKLLSISYAGOLGTVI 253

Db 149 FHPYTDFTV--WVQAVQGLAGLGVLPRAHGVKMFAPPLERSKLATTFITGSSVGMT 205

Qy 254 SLPISGIICYYNMWTVYFFPGTIGIFWFLMIWLVSDTPQKHRIISHYEKEYIISL 313

Db 206 GLPASAVYLSHFMSSTPFFVFGALGIWISILWIVSGSPETHGYISDDEKKQVTEKICD 265

Qy 314 ---QLSSQKSVPMVPIKSLPLMAIVAHFSYNMTFTLLTLPTYKKEILRFVQENGF 370

Db 266 VAVKMSLTTLPRMWTSTAVMAIILCTCRSMGFLLGNQLYMDVHLIDIKNSGF 325

Qy 371 LSSLPIYGLSWLCLMSGOADNLRAKVNFSTLCVRRISLIGMIGPAFLVAAGFIC-- 428

Db 326 ISIFPQGMCIYTLTISQSLDYLRSCKMSTPAVKSVNTRG-----FTVEAVMLGCLAF 379

Qy 429 ---DYSLAVALFTISTTLGFCSSGFSINHLDIAPSVAGILGINTPATIPGMVPIA 485

Db 380 FVRDPVIAVTCVIACGSGSVLSGFNVNHFIDIAFRHAPIIMGINGCAIAG-VGGMVT 438

Qy 486 KSLTPDNTVGEWQTVFYIAAIVNFGAIFFTLFAKEGVQNA 527

Db 439 NTVTYQNPDG-WKQVFLMAAIDIFGVIFPLFAKGVLPWA 479

RESULT 12

H84698
hypotheical protein At2g29650 [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004

[illegible]

A:Map position:5

Query March 22.2%; Score 629.5; DB 2; Length 501;
Best Local Similarity 31.6%; Pred. No. 5,1e-42;
Matches 153; Conservative 101; Mismatches 191; Indels 39; Gaps 14;

OY P L B G A R E A H A V Y C --- S A R Y N L A I L A F F G F F I Y V A L R V N L S V A L D V D N S T T L E D
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
D b P I S Q G I P - S O Q R P L T Y W M F P S M R F M T S I M L C F C F C V L M N S M G M A I V C M V N S S A T Y - D 63

OY 118 N R T S K A C E H S A P I K V H N N O T G K K Y Q M D A E F O G W L S G F P F Y I I T Q I R G V V A K I G S K 177
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
D b 64 N E T --- Y P E N A P L ----- L D W S D E Q G Y F S A F N A G L V M L F T G G - M A D K F N A K 109

OY 178 M L G F G I L G T V A L T L - F T - P I A A D L G V G P L I V L R A L E G L G E Y T F P M A H M A M S W A P L 234
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
D b 110 Y M I --- L V S V A I A S I A N F T L E M M A P I S Y V M A I F S F L V G P A D A L L Q P A M N S I T M F P T S 166

OY 235 E R S K L I S Y A G A O L G I V I S L P L S G I I C --- Y Y N M W T V Y F F F T G I F W F L M I W L V S 290
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
D b 167 E R S V A L G I A T G A Q R G I G T I I I P A G A L C S Q T E I F G G M P S I F Y L S G F I G V L F I F S Y I F L C A 226

OY 291 D T P Q K R I S H Y E X E Y I I L S L R N Q --- L S S Q K S V P W P I I L K S L M A I V A H F S Y N M F 346
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
D b 227 D K P S Q S C I S N E K F I T I S N Q S E D V G K R T E R K Y P M K N I L K S G A V M A S V I S L V C H E F L 286

OY 347 Y T L T L P L T Y W K E I L R F W O E N G F L S L P Y L G S W L C M I L S G A A D N L R - A K M F S T L C V 404
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
D b 287 M T L I M F L P S Y I A D V H N H T E N G I L S A L P Y S L W A K I G S S Y L N T W L O K N T W K K D T I C - 345

OY 405 R I F S L I G M I P A V F L V A G F I G C D Y S - L A V A F L T I S T T L G F C S G S G S I N H L D I A P S Y A 463
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
D b 346 - K V I N S I G I G I G F L L A F L A T F L D N E H A M W A V L F C L S M A S A G L H T P G C O L A L V S V A P A Y S 404

OY 464 G I L G I T T P A T I R G M W C P V J A K S L T P D N T Y G E M O T V Y I A A I N V F G A I F F T L F A K G E V 523
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
D b 405 G A V G G F T F F F A V S G I I H P I T T K M I V K D R T A E M W L V E Y I S T V I A I P I I I F N W G S T E V 464

OY 524 O N M A 527
|| ||
D b 465 O M W A 468

RESULT 14

T01534
hypocherical protein_A IG005110.nm - Arabidopsis thaliana
CISpecies: Arabidopsis thaliana (mouse-ear cress)
CDate: 19-Feb-1999 #sequence_revision 19-Feb-1999 #text_change 09-Jul-2004
CAccession: T01534
RAndrews, S.
submitted to the EMBL Data library, July 1997
A:Description: The sequence of A. thaliana IG005110.
AReference number: ZI4347
AAccession: T01534
AStatus: translated from GB/EMBL/DDBJ
AMolecule type: DNA
AResidue: 1-413 <AND>
ACross-references: UNIPROT:Q23065; UNIPARC:UPI0000049C57; EMBL:AFO13293; NID:g2252823
CExperimental source: cultivar Columbia
CGenetics:
A:Map position: 4
AIntrons: 79/3; 115/3; 128/3; 200/3; 288/3; 331/3; 353/3; 382/1
ANote: A.IG005110.nm
CSuperfamily: hexuronate transporter

Query Match 19.8% ; Score 562; DB 2; Length 413;
Best Local Smilarity 29.1%; Pred.No. 8.9e-37;
Matches 134; Conservative 61; Mismatches 136; Indels 130; Gaps 8;

OY 80 R Y N L A I L A F F G F F I Y V A L R V N L S V A L D V D N S T T L E D N R T S K A C E H S A P I K V H N N O T G 139
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
D b 61 R M V I L L C F S S F L L C N D R V N S I A I L P M ----S 90

```

OY      140 KKQWDAETQGMILGSPFYGYIIIOIPGVYASRIGGKMLGFGILGNVLTLPFIAD 199
Db      91 QSYNNSSATVGLIGSPFWMGYLTQIGIMDKFGK----- 128
OY      200 LGVGLIYLRALEGEGVTPPAMHAMWSSWAPLERSKLISYAGAQLGTVISLPLSG 259
Db      129 -----GVAMPANNMNLISKWIPVERSRSLATLVSGMYLDSVGLANSP 171
OY      260 IICYYMMNTVYFPFFGTGIGFWLLIMLVSDTPQKHKRISHYEKEYILLSLRQLSQK 319
Db      172 MLITKFGMPSPVFSYSGSLGSIWFLMLKFAVSSPDDDELSEEEKVILGSKPR-EBVT 230
OY      320 SVPMWPIILKSLPLMAIVVAHPSYNTFYTLITLLEPTY-----MKEL 361
Db      231 VIPWLLISKRPVWMLIIISHFCHNMGTFILITMPTTYNQARSSASVLSLFINICEBV 290
OY      362 RFNVQENGFSLSLPLTGLSMLCMILSGOADNLRAMNFTSLCVRRISFLIGMIGPAVPLV 421
Db      291 KFNLTRESLLCLLP-----MLTN-----AVPAN 313
OY      422 AAGFTGCDYSLAVAFLTISTLLGFCSSGFSINHLDIAPSYAGILGTTNPATPGWVG 481
Db      314 IGGWIA--DTLVSRGLSTINGSDAFSQGLSNHODIPRAYAGVLLGSLNTAGVLAGVGF 371
OY      482 FVIAKSLTPDNTVGEMQTVFYIAAANVGAIFLTFKGE 522
Db      372 TAATGYILQR--GSMDDVFKVAVALYLTIGLIVNMLFATGE 409

```

```

RESULT 15
T45634
hypothetical protein F13112.30 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004
C:Accession: T45634
R:Christians, N.; Robert, C.; Brottier, P.; Winkler, P.; Catcolico, L.; Artiguenave, F.; Sa
submitted to the Protein Sequence Database, November 1999
A:Reference number: Z23010
A:Accession: T45634
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-537 <CHO>
A:Cross-references: UNIPROT:Q9SD75; UNIPARC:UPI0000A2A98; EMBL:AL133292
A:Experimental source: cultivar Columbia; BAC clone F13112
C:Genetics:
A:Map position: 3
A:Introns: 35/2; 153/3; 229/2; 350/3; 399/3; 445/3; 457/1
A:Note: F13112.30

```

Query Match	18.7%	Score 531	DB 2	Length 537
Best Local Similarity	27.6%	Prod. No. 3,5e-34		
Matches	151	Conservative	86	Mismatches 195, Indels 116, Gaps 15

Qy	33	ASTPAHGVMS-----PVRLAAN-----DGEESTDRTLLPGAPR	69
		: : : : :	
Db	20	SSSSSLTGVIVSSFDWMRIPLPERDRRKLVLCTGRVYNSKFTNTVD----LCGSPR	75
		: : : : :	
Qy	70	AEAPVCCSAYN-----LALAFGCFIVYALRYNLVA	104
		: : : : :	
Db	76	HLRAVSCGDARTPEETAELTAQNPSEFITSERKVVAMLA-LALALCADRRVMSVA	134
		: : : : :	
Qy	105	LYDVYDSNTTLEDNKTACPBHSAPIKVHHNQTKYQWDAETQWILGSEFFGYIITQ	164
		: : : : :	
Db	135	IV-----PLSLSRQ-----WSKSFSGIYQSSSLMGYILSP	164
		: : : : :	
Qy	165	IPGGVASKIGKMLFGILGTAVLTFTPIAADLGVPILVLEALBGLGCVFFPMH	224
		: : : : :	
Db	165	IAGETLYDRYGGKVMANGVALMSLATLTLPFAADSSLMALLAARAMGVAEVALLPCMN	224
		: : : : :	
Qy	225	AMMSMAPEPERSKLLISVGAQLGVYISLPLSGITCYNNMYVVFYFFGTIGLFWLL	284
		: : : : :	
Db	225	NMVARWPEPTERSRAVGIAMGFQDGNVVGMLSLIMSOGIYGPVPIFGISGLTWLLAV	284
		: : : : :	

[illegible]

Search completed: June 7, 2006, 05:43:41
Job time : 44 secs

Job time : 44 secs

THIS PAGE BLANK (USPTO)


```

Db      63 RGPQSVFLAEALGASWSLIMFKYATDBKSTASGVGSVLFPVKKIDITHNKK----- 114
      314 QLSQSKSVPMVPLIKSLPLMAIVAAHFSYNTFTLTLLPTYNKELIRFVQENGFLSS 373
      115 PLSAK--IPWVKILTSFPWAIIVNNFTFHALVLMNMVLTPTYPELGQLSLQDMGSSKM 172
      374 LPIYGLWLCMLISGOADNLRKKNFSTLCVRIFSLIGMIGPAVFLVAGFICDYSLA 433
      173 MPYLMNLFNSNIGVADVILTRILSVTKRKFLNVG-----FLVA-----SLA 218
      434 VAFLTISTVLGG--FCGS-----GFSINHLDIAPSYAGILGINTPATIPGMV 480
      219 LVIIIPRFTSGAVFCSSVALGFLALGRAGFAVNMHDIAPRYAGIVMGSVSTAGTLAGIV 278
      481 GPVI-----AKSLTPDNVTGE--MOTVFYIAAIAINFGAFTFLPAKGE 522
      279 GVDLTGKLEAKKANSDLSSESRAVFSIPGFLCIFSSVFILFSTGE 328

```

RESULT 5

```

US-10-471-571A-3576
; Sequence 3576, Application US/10471571A
; Publication No. US20060115490A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026927NO
; CURRENT APPLICATION NUMBER: US/10/471.571A
; PRIOR FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: GB-0107661.1
; NUMBER OF SEQ. ID NOS: 5642
; SOFTWARE: SeqWin99, version 1.03
; SEQ. ID NO 3576
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (1)..(425)
; OTHER INFORMATION: hypothetical protein
US-10-471-571A-3576

```

Query Match 13.5%; Score 381.5; DB 6; Length 425;
 Best Local Similarity 26.0%; Pred. No. 4.8e-24;
 Matches 108; Conservative 81; Mismatches 177; Indels 49; Gaps 10;

```

      78 SARYNLAIIAIFGFFIYVALRVNLVALVDMVDSNTTLEDNRKTSKACPEHSAPIKVAHQ 137
      7 NVKMFALAFIFGIVAIYMBRANISYIAKQMD-----DLGMYK----- 45
      138 TGGKYQMDAETQCMILGSPF--YGYITQIDPGYVASKIGGKMLIGFISLITAVTLTFTPI 196
      46 -----PQGLLASFFSLGAYALMQVPSGMLAEKFGGRKMTITIALVWMSAFTILGTM 95
      197 AADLGVPLIVLALBGLGEGVTPRAMHAMSSAPRLERSKLIISYAGAGQGTIVSLP 256
      96 IKNHGL--IYLVFPLFGVGAAPMIPSNVANSFESKNEKGRASSALLAGSFGFVLAPI 153
      257 LSGIICYYMMTVVFFPFGTIGIFWFLMWLVSDTPQKHRIISHYEKEYILSLRNQLS 316
      154 VTIAIVAAFMQAVFYFGAVGILMAVLMIAKDPDEGRMVNVEAKRFIMENRDI VAT 213
      317 SQKSVPMVPLIKSLPLMAIVAAHFSYNTFTLTLLPTYNKELIRFVQENGFLSSLPY 376
      214 EKSSPPNDFPKKFSFPAIAIOYFVVQFIITLFIWLPYLTLEVFHNFKEMS--ISLPLW 272
      377 LGSMLCMLISGOADNLRKKNFSTLCVRIFSLIGMIGPAVFLVAGFICDYSL 432
      273 LMFPLILSGAISDRVLGGRSKF-----VARGVIAIAGFIYFAVSIITFAVRGTG---NL 324
      433 AVAFLTISTVLGG--FCGS--INHLDIAPSYAGILGINTPATIPGMVGPVIA 485

```

```

Db      325 YVSIFWLSLIGGIGISMGWSMAATDLGRNFSGTVSGMMNLMGNIGALISPLIA 379

```

RESULT 6

```

US-10-953-349-14672
; Sequence 14672, Application US/10953349
; Publication No. US20060107345A1
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, NICKOLAI et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-1579PUS2
; CURRENT APPLICATION NUMBER: US/10/953.349
; PRIOR FILING DATE: 2004-09-30
; NUMBER OF SEQ. ID NOS: 40252
; SOFTWARE: Patent version 3.3
; SEQ. ID NO 14672
; LENGTH: 288
; TYPE: PRT
; ORGANISM: Glycine max
US-10-953-349-14672

```

Query Match 13.1%; Score 371; DB 6; Length 288;
 Best Local Similarity 32.2%; Pred. No. 2.2e-23;
 Matches 100; Conservative 51; Mismatches 100; Indels 60; Gaps 11;

```

      244 YAGAGLCTVLSLPLSGIICYYMMTVVFFPFGTIGIFWFLMWLVSDTPQK----- 295
      2 YLGAALGMLL-LP---TLVFRGPQSVFLAEALGASWSLIMFKYATDPKSTASGVGSBV 57
      296 ---HKRISHEKEYILSLRNQLSOKSVPMVPLIKSLPLMAIVAAHFSYNTFTLTLL 352
      58 LPVKKIKIDITHNKK-----PLSAK--IPWVKILTSFPWAIIVNNFTFHALVLMNMW 107
      353 LPTYMKELIRFVQENGFLSSLPYLSGLWLCMLISGOADNLRKKNFSTLCVRIFSLIG 412
      108 LPTYFELGLQSLQDMGSSKMPYLMNLFNSNIGVADVILTRILSVTKRKFLNVG 167
      413 MIGPAVFLVAGFICDYSIAVAFLTISTVLGG--FCSS-----GFSINHLDA 459
      168 -----FLVA-----SLALVIIPRFTSGAVFCSSVALGFLALGRAGFAVNMHDI 213
      460 PSYAGILGINTPATIPGMVGPVI-----AKSLTPDNVTGE--MOTVFYIAAIAVFG 511
      214 PRYAGIVMGVSNTAGTLAGIVGVDLTGKLEAKKANSDLSSESRAVFSIRGFLCIF 273
      512 AIFTFPAKGE 522
      274 SPVFLFSTGE 284

```

RESULT 7

```

US-10-471-571A-1940
; Sequence 1940, Application US/10471571A
; Publication No. US20060115490A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026927NO
; CURRENT APPLICATION NUMBER: US/10/471.571A
; PRIOR FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: GB-0107661.1
; NUMBER OF SEQ. ID NOS: 5642
; SOFTWARE: SeqWin99, version 1.03
; SEQ. ID NO 1940
; LENGTH: 452
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (1)..(452)

```

```
; OTHER INFORMATION: glycerol-3-phosphate
US-10-471-571A-1940

Query Match      7.9%; Score 223.5; DB 6; Length 452;
Best Local Similarity 22.9%; Pred. No. 4.4e-11;
Matches 109; Conservative 75; Mismatches 191; Indels 101; Gaps 19;

QY      83 LAILAFGFPPIVYA---LRNLVALVDMVDSNTLLEDRKTKACPEHSAPIKVHNRQ 138
DB      25 LRLQVFLGIFIGAGYVLLRKNFSLAM-----PALQEGCFKA-----62

QY      139 GKXQVDAETQGIISFFGYIITQIPGVYASKIGKMLLGFGLGTAVLTP---FTF 195
DB      63 -----ELGPAISAVSIAVGFSPKFGTSDRSNARIFVLGLALTIIVLLMGFVP 113

QY      196 IADLGVPLIVLRALEGLGEGVTPPAMHAMSSMAPPLERSKLLISYGAQLGTVISL 255
DB      114 FPTS-GIGIMFVLLFLNGMFGMGWPPSGRVLVHMSVSERGSKTALMNVANHVGGINA 172

QY      256 PLS-GI-ICYNMNTY-----VFYFPGTIGIFWFLMTWLVSDTPQKH--KRISYK 304
DB      173 PIAAMGITTTFAPINFGYLGKFBEGFTYPPALAILIISAISYVLLRDPQSGLPPIETIKN 232

QY      305 EYLS---SLRNQSSQKSVPMVPIKSLPLMAIVAHFSYNMTFYTLTLPTYMKEL 361
DB      233 DFATSDKKTLETLELTI-KEILFKYVLLNNKVMALAPANITFYVRIGVLDMAVYLSSEK 231

QY      362 RFNVQENG---FP---SSLPYLGSLCMLISGQAADNLRAKNPSTLCVRRIFSLIGMIG 415
DB      292 HPDLKASGMAYFYEWAGIP--GTLGCGYISDKLFKGRGPAGFFFWLGVTVFLLIYMLN 349

QY      416 PA-----VFLVAAGFIGDYSLAVAFLITSTLGFSCSGSINHLDIAP--SYAGILL 467
DB      350 PGNAMLDMVSLAIGFLIYGPVLI-----GLQALDVPKKAQTAA 392

QY      468 GINTFATIPGMV-----GPVIAKSLTPNTVGEMQTVFYIAAINVFAGIFTL 517
DB      393 GLTGLFGYLFGAVMANIVLGAIVDKF-----GMDVGFIILTLAISVFAMLSLIL 440

RESULT 8
US-10-471-571A-2552
; Sequence 2552, Application US/10471571A
; Publication No. US20060115490A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026927MO
; CURRENT APPLICATION NUMBER: US/10/471,571A
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: GB-0107661.1
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 5642
; SOFTWARE: SeqWin99, version 1.03
; SEQ ID NO 2552
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (1)..(465)
; OTHER INFORMATION: hexose phosphate transport
US-10-471-571A-2552

Query Match      5.6%; Score 157.5; DB 6; Length 465;
Best Local Similarity 20.2%; Pred. No. 1.1e-05;
Matches 99; Conservative 81; Mismatches 178; Indels 133; Gaps 23;
```

```
DB      72 -----LGYIGLAFSTIYGLGKTLIGYFVDGRNTRKRIISFLILSAITVLIMGFVLSYFGS 126
QY      202 -VGPLIVPALSGLGEGVTPPAMHAMSSMAPPLERSKLLISYGAQLGTVISLPLSGI 260
DB      127 VNGLLIVLWGLNGVFGQSVGGPASYSFISRMABRTKRGRLGFWNTSHNIGAI---GGV 183
QY      261 ICYNMNTYVFYFPGTIGIFWF-----LLMTWLVSDTPQKH--KRISYKEYIIS 309
DB      184 ALMGAN---VFPHGNVIGFIFPSYALLIGATLFIQDDDBELQMNABEIEWEPV-- 238
QY      310 SLRNQSSQKSVPMV---VPIKSLPLMAIVAH--FSY-----NMTFYTLTLPTYM 357
DB      239 -DKENIDSGQMTKWEIFKKYIILGNPVYITLCSNVFVYIVRIGIDMA-----PLVY 289
QY      358 KEILRFNVQENGFLSLPY--LGSMLCMLISGQAADNLRAKNPSTLCVRRIFSLIGMIG 415
DB      290 SEHLHFS--KGDVAVNTIFYEIGALVASLLMGVSDLLG-----RRAIVAIG--- 335
QY      416 PAVFLVAAGFIGDYSLAVAFLITSTLGFSCSGSINHLDIAPSYAGILLGIT----- 470
DB      336 -CMEMIT--FVLFYTNATSWMVNISL-----FALGALLFGPQ---LLIGVSLTGRV 382
QY      471 --NTFATIPGMVGPV-----IAKSLTPDN-----TVGEMQTVFYIAAIN 508
DB      383 PKAIVSIVANGMTGSFAVYLFQDSMAKVGLAIAIDPTENGINIGYTLGWTDFVIFYVAL 442
QY      509 VEGALFPTLFA 519
DB      443 FLGMILLGIYA 453
```

```
RESULT 9
US-10-471-571A-3466
; Sequence 3466, Application US/10471571A
; Publication No. US20060115490A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026927MO
; CURRENT APPLICATION NUMBER: US/10/471,571A
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: GB-0107661.1
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 5642
; SOFTWARE: SeqWin99, version 1.03
; SEQ ID NO 3466
; LENGTH: 388
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (1)..(388)
; OTHER INFORMATION: quinolone resistance
US-10-471-571A-3466

Query Match      5.4%; Score 152.5; DB 6; Length 388;
Best Local Similarity 19.7%; Pred. No. 2.3e-05;
Matches 76; Conservative 75; Mismatches 158; Indels 77; Gaps 13;

QY      150 GWILGSEFFGYIITQIPGVYASKIGKMLLGFGLGTAVLTLFTPIADLGVGPLIVLR 209
DB      41 GLVVAAPALSSQMTISPPGGLADKLKGKILICIGLIFSVSEFMFAVGHNFV--LMLSR 98
QY      210 ALBGLGEGVTPPAMHAMSSMAPPLERSKLLISYGAQLGTVISLPLSGIICYNMNTY 269
DB      99 VIGGMSAGMVPVGTGLIADISPSHOKAKNFGYMSAIIINSGLTGPQIGFMA--EVSIRM 157
QY      270 VFYFPGTIGIFWFLMTWLVSDTPQKHKRISYKEYEYIISLRNQLSSQKSVPMVPIKLS 329
DB      158 PFPFAGALGLAFIATVSLVLIHD--PKKSTTSGFQ-----LEPQLITK--INMKVFITP 207
QY      330 LPLMAIVAHFSYNMTFYTLTL-LPTYMKELRFNVQENG-----FLS 372
```

Db 208 VILTLVLSFGLSAETIYSLYTADKVNYSPKDISIAITGCGIFGALFQIYFEEDKMKYFS 267
Qy 373 SLPLIG-----SWFCMLISGOADNLRAKNFSTLCVARIISLIGMIGPAFLVAAGRTG 427
Db 268 ELFTIAMSLSYVVLILPLVAND---YWSIMLSIF-VVITGDMIRPAL----- 313
Qy 428 CDVSLAAVFLTISTTLGFCSSGSINHLDIAPSVAGILGITTTFATIPGMVGVIAKS 487
Db 314 -----INYS---NIGERQSGAGLNGFTSMGNFIFGLIAGA 349
Qy 488 LTPDNTVGEQWTFVYIAAIVFCAI 513
Db 350 LFDVHI---EAPIYMAIGVSLAGV 371

RESULT 10
US-10-471-571A-5276
; Sequence 5276, Application US/10471571A
; Publication No. US20060115490A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026927WO
; CURRENT APPLICATION NUMBER: US/10/471,571A
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: GB-0107661.1
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 5642
; SOFTWARE: Seqwin99, version 1.03
; SEQ ID NO 5276
; LENGTH: 466
; TYPE: PRP
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (1)..(466)
; OTHER INFORMATION: hypothetical protein
US-10-471-571A-5276

Query Match 5.3%; Score 151; DB 6; Length 466;
Best Local Similarity 22.0%; Pred. No. 3.8e-05;
Matches 93; Conservative 66; Mismatches 154; Indels 110; Gaps 17;

Qy 153 LGSFPGYITTOIPGCVVASKIGGKMLGFGILGTAV---LTLPPIAADLGVGPLVL 209
Db 55 LSLPLFAGLFI--VGAGDVADKFGKRVKITVVGGLINNVGSLIITPLPAFLITG----R 107
Qy 210 ALBGLGEGVTFPAMHAMSSWAPPLERSKLLSISVAGAOIGTVISLPLSGIICYNNMTY 269
Db 108 IIGLSAACIMPSLTAIINEYIGTRORALSYWSISGSGSICLIFGLMATYIGMR 167
Qy 270 VFYFPGTIGIFWFLMIWLVSDTPQKH---KRISHYEKEYILSL-----RNQLS 316
Db 168 IF-----VVSILTLTLMAYLIKHAPEYKAEPKGMKAKKFDVGLVILVMTLSLVII 223
Qy 317 SOKS-----VPMV-----PIKSLPLMAIVAAHSYN 343
Db 224 TQTSHGVLSPLLGLIVPFCISLVGFVYENKIKHPLVDPSIFKNRGSGATISNPLN 283
Qy 344 WTFYTLTLTLPTYMKEILRFNVOENGFLSLPYLGSWLCMILSGOADNLRAKNFSTLC 403
Db 284 GVGAGALIVINTYYQQLGNSGOTGYI-SLTYIITVLSMIRVGEKI--LSQHGRKRPPL 340
Qy 404 VRIIFSLIGMI-----GPAVFLVAAGFICGDYSLAAVFLTISTTLGFCSSGSFSINHL 457
Db 341 LSGGFVYIGLILSLFLPLPEWVYIISIVG-----YLLFGTGLG-----LY 381
Qy 458 IASVYAGILGINTTFATIPGMGVIAKSLTPDNTVGEQWTFVYIAAI-NVSG-AIFF 515
Db 382 ATPS-----TDT-----AVASAPDUSGVASGVYKMASSLSGNAGVAVSG 421
Qy 516 TLF 518

Db 422 TVY 424

RESULT 11
US-10-471-571A-4604
; Sequence 4604, Application US/10471571A
; Publication No. US20060115490A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026927WO
; CURRENT APPLICATION NUMBER: US/10/471,571A
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: GB-0107661.1
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 5642
; SOFTWARE: Seqwin99, version 1.03
; SEQ ID NO 4604
; LENGTH: 462
; TYPE: PRP
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (1)..(462)
; OTHER INFORMATION: hypothetical protein
US-10-471-571A-4604

Query Match 5.0%; Score 143; DB 6; Length 462;
Best Local Similarity 21.8%; Pred. No. 0.00017;
Matches 94; Conservative 65; Mismatches 173; Indels 100; Gaps 19;

Qy 153 LGSFPGYITTOIPGCVVASKIGGKMLGFG---LIGTAVTLTLPPIAADLGVGPLVL 208
Db 53 LTLFSGMEV--VGAGLADKIGKRVKNTNIGLISITGSA-LITTLVPLALITG----- 104
Qy 209 RALEGLGEGVTFPAMHAMSSWAPPLERSKLLSISVAGAOIGTVISLPLSGIICYNNMT 268
Db 105 RVIGVSAACIMPSLTAIMKTYOGAERQALSYWSISGSGSICLIFGAVATTMGWR 164
Qy 269 VFYFPGTIGIFWFLMIWLVSDTPQKHKRISHYEK----- 304
Db 165 WIFIF---SIIYAVLSMLIKGTPERKSEITNTHKFDVAGLIVLVMLSLNVITKGA 220
Qy 305 -----EYILSLRNQLSQKSY--PMV--PIKSLPLMAIVAAHSYNMTFY 347
Db 221 ALGYSLMFGLAIVVAFPIPLNVEKKVDNPLIDFKLENNRYTGATTISNPLN-GFA 279
Qy 348 TLTTLTPTYMKEILRFNVOENGFLSLPYLGSWLCMILSGOADNLRAKN-----FST 401
Db 280 GTLIVANTFPVQGLGYALQAGYL-STTYIIMVLMIRVG---EKLQKSGSKRPMILGT 335
Qy 402 LCYRRIIFSLIGMI--PAVFLVAAGFICGDYSLAAVFLTISTTLGFCSSGSFSINHLDIAP 460
Db 336 FIWVIGIALISLVLPQIFVY---ISC---VVGYLCPGLGIVATPS-----TDTAI 382
Qy 461 SYA-----GILLGITNTFAIIPG-----MGEVIAKSLTPDNTVGEQWTFVYIAA 505
Db 383 SNAPLDKRVGASGYKMASSLSGAFGVALISGAVYAGVAATSIHTGMALMNVNL---M 439
Qy 506 AINVFCAIFFTL 517
Db 440 GIMAFALIFAI 451

RESULT 12
US-10-471-571A-4394
; Sequence 4394, Application US/10471571A
; Publication No. US20060115490A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026927WO

```

; CURRENT APPLICATION NUMBER: US/10/471.571A
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: GB-0107661.1
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 5642
; SOFTWARE: SeqWIn99, version 1.03
; SEQ ID NO 4394
; LENGTH: 643
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (1)-(643)
; OTHER INFORMATION: hypothetical protein
US-10-471-571A-4394

Query Match
Best Local Similarity 5.0%; Score 142.5; DB 6; Length 643;
Matches 11; Conservative 72; Mismatches 186; Indels 171; Gaps 24;

QY 79 ARYNLAIAFFGFPIYVALRVNLVAVDNDSTTIEDNRTSKACEHSAPIRVHINOT 138
DB 137 SRGKIIAALIFGFIALLNQTLNLVALPKI--NT----- 168
QY 139 GKRYQMDAETQGMILGSEFY--GYITPQIGYVASKIGKMLGFGILGTAVALTLPPI 196
DB 169 --EPNISASTGQHLMTGFMVLVNGILPIT--AIFPKYSYRKLFLVALVLTIGSLICAI 224
QY 197 AADLVGPFILVLRALBGLGEGVTFPPAMAMSSWAPLERSKLISISYAGAQGTV--- 252
DB 225 SNNFPI--MMVGRVLQAIAGAVLMPGLGSIYITITYPEKR-----GAMGTGMIAM 273
QY 253 -----ISLPSGIIICYMMNTVYFFGDTGIF-----WELIMVLV--SPTP--- 293
DB 274 ILAPALGPTLSGYIVQNVHNMVFMFGMFIIGIILAILGFWFKLYQYTTNPKADIPGIIF 333
QY 294 -----QKKRISHYEKE-----YILSLRNQLSOKSVPWVPIKSLP 331
DB 334 STIGFGLALYGFSEAGKMGSVIEITMFAIGIIFILFVIRELRMSPMLEVLKFP 393
QY 332 LMAIVAHFSYNTFTYLLTLLPTYMKEILRNVQENGLFSSLPYLGSLCMIL---SQQ 388
DB 394 FTLTITINMVMVLSYGMILLPIYQLNLRGFSALDGLL-LFP--GSLIMGLGPPAGK 450
QY 389 AANLRAK-----WNFSTL-----CVRIFSLIGIGPAY--- 418
DB 451 LDDTIGLKPLAIFGIAMVTYATWELTYLNDPTPYMTIMGIYVLSFGMAFIEMMTVAI 510
QY 419 -----FL-----VAAGFICDYSALAVLTISTT--LGFCSSGFSIN--- 454
DB 511 NALPGRASHGNAFLNMRQLASIGR--AIVTVMTTQTTQHLSAFGBELDKTNPVVD 568
QY 455 HL-DIAPSVAG-----ILLGITNTFATIPGMVGPVIAKSLTDPNTVGEWQVFIYAAI 507
DB 569 HMBELASQYGGGEGAMKVLLOFVVKLATVEGINDAFIV-----ATIPSIIALI 616

RESULT 13
US-10-471-571A-2776
; Sequence 2776, Application US/10471571A
; Publication No. US20060115490A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026927NO
; CURRENT APPLICATION NUMBER: US/10/471.571A
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: GB-0107661.1
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 5642
; SOFTWARE: SeqWIn99, version 1.03
; SEQ ID NO 2776
; LENGTH: 418

```

```

; TYPE: PRT
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (1)-(418)
; OTHER INFORMATION: hypothetical protein
US-10-471-571A-2776

Query Match
Best Local Similarity 4.9%; Score 139.5; DB 6; Length 418;
Matches 78; Conservative 70; Mismatches 169; Indels 65; Gaps 16;

QY 168 GYVASKIGKMLGFGILGTAVALTLF-----TPAADLVGPIVLRALBGLGEGVTFPA 222
DB 70 GKLGDIRSKMMVLRALLGLAVCLFLMALCTTFL-----QVLVRLQLGFLGCVV-DA 121
QY 223 MHAMSSWAPLERSKLISISYAGAQGTIVISLPSGIIICYMMNTVYFFGDTGIFNF 282
DB 122 SSAPASAEAPAEEDRGVILGRLOSVSAGSLVGPLIGVASIIIGFALIM--SIIVTF 178
QY 283 LLMIV-----LVSDTPQKKRISHYEKEYILSLRNQLSOKSVPWVPIKSLPLMAIYVA 338
DB 179 IVCLFGALKIETTHMPKSGTPINIKG--IRRSFQCLCTQOTC-----RFIIV 225
QY 339 HFSYNTFTYLLTLLPTYMKEILRNVQEN--GFLSSLPYLGSLCMILSGAADN--L 393
DB 226 GVLANFAMGMFLALSPASSVNHATAIDRSVIGFLQSAFWTASIIAPLMGRFNDKSYV 285
QY 394 RAKMNSTLCVRKRIIFSLIGIGPAVLVAAGFI-GCDYSL--AVAFLTISTLGGFCS 449
DB 286 KSVYIATTAICGSAILOGATNIEFLMAARILOGLTYSALQSVFVNVNA-----CHQ 340
QY 450 GFGINHLDIAPSYAGIILGTTNTFATIPGMVGPVIAKSLTDPNTVGEWQVFIYAAIN 509
DB 341 -----QAKGFVGTGNSMLYVGQILGSLGSAITSYTPA---TTFIVMGVFA 386
QY 510 FGAIFFTLPAKGEVQWALNDH 531
DB 387 VSSLFLIC--STITN-QINDH 404

RESULT 14
US-10-471-571A-5294
; Sequence 5294, Application US/10471571A
; Publication No. US20060115490A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026927NO
; CURRENT APPLICATION NUMBER: US/10/471.571A
; CURRENT FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: GB-0107661.1
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 5642
; SOFTWARE: SeqWIn99, version 1.03
; SEQ ID NO 5294
; LENGTH: 402
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (1)-(402)
; OTHER INFORMATION: hypothetical protein
US-10-471-571A-5294

Query Match
Best Local Similarity 4.8%; Score 135; DB 6; Length 402;
Matches 92; Conservative 77; Mismatches 182; Indels 120; Gaps 20;

QY 81 YNLALIAFFGFPIYVALRVNLVAVDND-----SNTLEDNRTSKACEHSAPIKVA 134
DB 6 FKLSIMKLTFTI-----LSIFIVGMVEMVAGIMNLSQDLHSEA----- 46

```

```

QY 135 HNOTGKKYQMDAETQGMILGSFFYGIITQIPGGVASKIGCKM-----LLGFGIL 185
DB 47 -----VGGOLVMTKALTFAICGPILVLTNRSSRPVLLMTLILF-II 88
QY 186 GTAVLTFTPIADLVGVPILVLRLBGLGEGVTFPPAMHAMSSWAPPLERSKLLSISYA 245
DB 89 GNGIIA----VAPNFSI--LVVGRILISSAAALIIYKVALITAMLSAPKNGKMGILVYT 142
QY 246 GAQGTYSILPLSGIICYYMMTYVFYFPGTIGF-WLLMIMLVSDP-----PQKHRI 299
DB 143 GFSGANFVGVPVIGTIVIGDLVGMRTYFLELIIVSIIIVGLMMIYLPKQEIQRGPVNHETP 202
QY 300 SHYEKEYILSLRQLSQKSVPMVPPIKSLPLMAIVAHFSYMTFTYLLTLPTWKKE 359
DB 203 SH-----EHHVTSKIIRP-AEVAKYLIITFVLIANSVFVFNPIILNSGHMS 251
QY 360 ILRFNVQENGFLSLPYLGSWLCMLISQAADNLRK-WNESTLCVRRIEFLI-GMIGPA 417
DB 252 FVSLALVNG-----IAGVIGTSLGIFSDKITSKMLMSISIFVMMMLMNLILPG 304
QY 418 VFLLVAGPIGCDYSLAVAFLLISTTLGFCSSGFSINHLD-----IAPSYAGILL 467
DB 305 SGLLALGF-----IWNIMQWSTNPVQSGV-IQHVBDTQVMSMNSSLNAGIGV 355
QY 468 GITVTFATIPGMVGPVIAKSLTPDNTVGEWQTFVYIAAIVFGAI-FTTL 517
DB 356 G-----GIIGGLVMTNHSV-----QAITTSALIGALGIIVFTL 390

```

```

RESULT 15
US-10-471-571A-3330
; Sequence 3330, Application US/10471571A
; Publication No. US20060115490A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026927W0
; CURRENT APPLICATION NUMBER: US/10/471,571A
; PRIOR FILING DATE: 2003-09-12
; PRIOR APPLICATION NUMBER: GB-0107661.1
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 5642
; SOFTWARE: SeqwIn99, version 1.03
; SEQ ID NO 3330
; LENGTH: 470
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (1)_(470)
; OTHER INFORMATION: proline/betaine transporter
US-10-471-571A-3330

```

```

Query Match          4.8%; Score 135; DB 6; Length 470;
Best Local Similarity 21.3%; Pred. No. 0.00079;
Matches 93; Conservative 74; Mismatches 172; Indels 98; Gaps 21;

```

```

QY 139 GKXKQW-----DAETQGMILGSFF-----YGIITQIPG---GYVAS 172
DB 30 GNMMEWDFGVYAATYATYIGANFSPVENDIRQMLTFAALALAFLLRPIGVVFGIIGD 89
QY 173 KIGGKMLGFGILGTAVLT---FTPIADLV-GPLIVL--RALEGLEGVTFPPAMHAM 226
DB 90 KYGRKVVLTSTIIIMAFSTLTIGLLPSYDQIGLWAPILLARVLQGFSTGGEYAGANTY 149
QY 227 WSSWAPPLERSKLLSISYAGNQLGTVISLPLSGIICYYMMTYVFYF-----FG--- 275
DB 150 VAESSPDKRRNSLGS-----GLEIGT-----LSGYIASIMAVLTFPLTDEQMASFGWRI 200
QY 276 --TIGIFWFLIMWL---VSDTPQKHRIKISHYEKEYILSLRNQSSQ---KSVPMVPIIL 327
DB 201 PFLIGLFLGLFGLYLRKRLRESF-----VFENDVATQPERDININFLQII 244

```

```

QY 328 KSL--PLMAIVAHFSYNTFTYLLTLPTWKKEILRFNVQENGFLSLPYLGSWLCMIL 385
DB 245 RPYKDI FVCFVANVFENNVTNNMTAYLPYIEGVIKLDNTTTSVLTCWAMIMPLAM 304
QY 386 SGQADNLRKWNNESTLCVRRIEFLIGMIGPAFLVAGFIGCDYSLAVAFLLISTTLGG 445
DB 305 FGKLADRIKGEK-----KVF-LIGTGGLTFSLIAFMILHSQSFVV--IVIGIFILG 352
QY 446 FCSGFSINHLDIAPS--YAGI---LIGIT-NRFATIPGMVGPVIAKSLTPDNTVGEWQT 499
DB 353 FFLSTYEATMPSLPTMFYSHIRYRTLSVFNISVSIFGGTTPLVAVTLV--TKTGDPILA 410
QY 500 VFYIAAIVNFGAIFFT 516
DB 411 PAYLTAISVIGFLVIT 427

```

```

Search completed: June 7, 2006, 05:48:14
Job time : 18 secs

```

THIS PAGE BLANK (USPTO)

Protein Sequence Searches - February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension **.rup**) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (UniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: June 7, 2006, 05:34:31 ; Search time 301 Seconds
(without alignments)
1647.204 Million cell updates/sec

Title: US-10-823-506-8
Perfect score: 2836
Sequence: 1 MAGAMTPRPVQPARPGGF.....LFAKGEVQNALNDHGHHRH 536

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt-7.2:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2621	92.4	495	1 S17A5_HUMAN	G9nar2z homo sapien
2	2621	92.4	495	2 O5S276_HUMAN	Q3n276 homo sapien
3	2329	82.1	495	1 S17A5_SHEEP	Q3m2d1 ovis aries
4	2298	81.0	495	1 S17A5_MOUSE	Q3bnr2 mus musculu
5	2298	81.0	495	2 O3TE25_MOUSE	Q3te25 mus musculu
6	2298	81.0	495	2 O5Q0U0_RAT	Q3q0u0 rattus norv
7	2072.5	73.1	484	2 O5ZL94_CHICK	Q3z194 gallus gall
8	1582.5	55.8	490	2 O4SGF2_TETNG	Q4sgf2 tetradon n
9	1051	37.1	473	2 O7OS80_ANOGA	Q7os80 anopheles g
10	996.5	35.1	559	2 O9V167_DROME	Q9v167 drosophila
11	992	35.0	502	2 O9VDM0_DROME	Q9vdm0 drosophila
12	991	34.9	469	2 O7O579_ANOGA	Q7o579 anopheles g
13	987	34.8	493	1 YLID2_CAEEL	O03567 caenorhabdi
14	970	34.2	582	2 O920B7_MOUSE	O920b7 mus musculu
15	970	34.2	582	2 O9J112_RAT	Q9j112 rattus norv
16	966.5	34.1	584	2 O8A47_BRARE	Q8a47 brachydanio
17	965.5	34.0	582	2 O9P2U8_HUMAN	Q9p2u8 homo sapien
18	965	33.9	582	2 O8BL57_MOUSE	Q8bl57 mus musculu
19	960.5	33.9	397	2 O4SY72_TETNG	Q4sy72 tetradon n
20	957.5	33.8	584	2 O5M818_BRARE	Q5m818 brachydanio
21	956	33.7	504	2 O4S4P9_TETNG	Q4s4p9 tetradon n
22	955	33.7	589	2 O8NDX2_HUMAN	Q8ndx2 homo sapien
23	954.5	33.7	588	2 O7TSF2_RAT	Q7tsf2 rattus norv
24	954.5	33.7	588	2 O8K1Q1_RAT	Q8k1q1 rattus norv
25	946.5	33.4	474	2 O2PJ91_CAEEL	Q2pj91 caenorhabdi
26	946.5	33.4	479	2 O23514_CAEEL	Q23514 caenorhabdi
27	945	33.3	601	2 O8BFU8_MOUSE	Q8bfu8 mus musculu
28	944.5	33.3	587	2 O5M817_BRARE	Q5m817 brachydanio
29	941	33.2	560	2 O9P2U7_HUMAN	Q9p2u7 homo sapien
30	939.5	33.1	576	2 O6INC8_XENLA	Q6inc8 xenopus lae
31	939	33.1	474	2 O623G6_CAEER	O623g6 caenorhabdi

32	938	33.1	573	2 O4S7G2_TETNG	Q4s7g2 tetradon n
33	936	33.0	560	2 O3TXX4_MOUSE	O3txx4 mus musculu
34	936	33.0	560	2 O62613_RAT	O62613 rattus norv
35	932	32.9	560	2 O6PCD0_HUMAN	O6pcd0 homo sapien
36	928.5	32.7	529	1 P1CO_DROME	Q9v785 drosophila
37	927.5	32.7	479	2 O3ZLF0_BOVIN	Q3z1f0 bos taurus
38	917	32.3	483	1 P1CO_DROAN	O61369 drosophila
39	913.5	32.2	576	2 O715L3_XENLA	O715l3 xenopus lae
40	910.5	32.1	461	2 O7PV86_ANOGA	O7pv86 anopheles g
41	910.5	32.1	502	2 O61IR3_CAEER	O61ir3 caenorhabdi
42	906	31.9	478	2 O5S2A1_MOUSE	Q5s2a1 mus musculu
43	894.5	31.5	472	2 O7Q367_ANOGA	Q7q367 anopheles g
44	889.5	31.4	508	2 O7O1S5_ANOGA	O7o1s5 anopheles g
45	876	30.9	497	2 O9Y2C5_HUMAN	O9y2c5 homo sapien

ALIGNMENTS

RESULT 1	ID	S17A5_HUMAN	STANDARD;	PRT;	495 AA.
AC	Q5NRAZ	Q8NERS	Q9UGH0		
DT	07-JUN-2004	Integrated into UniProtKB/Swiss-Prot.			
DT	07-JUN-2004	sequence version 2.			
DT	07-FEB-2006	entry version 27.			
DE	Sialin (Solute carrier family 17 member 5) (Sodium/sialic acid cotransporter) (AST) (Membrane glycoprotein HP59).				
GN	Name=SLC17A5;				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1), FUNCTION, AND TISSUE SPECIFICITY.				
RX	PubMed=11751519;				
RA	Fu C., Bardhan S., Cetateanu N.D., Wani B.D., Wang Y., Yan H.-P., Shi E., Carter C., Venkov C., Yakes F.M., Page D.L., Lloyd R.S., Mernagh R.L., Hellerqvist C.G.;				
RT	"Identification of a novel membrane protein, HP59, with therapeutic potential as a target of tumor angiogenesis.";				
RL	Clin. Cancer Res. 7:4182-4194(2001).				
RN	[2]				
RP	NUCLEOTIDE SEQUENCE [MRNA] (ISOFORM 1), FUNCTION, TISSUE SPECIFICITY, AND VARIANT SD CYS-39, AND VARIANTS ISSD 268-SER-ASN-272 DEL; ARG-183 AND ARG-334.				
RX	MEDLINE=20047778; PubMed=10581036; DOI=10.1038/70585;				
RA	Verheijen F.W., Verbeek E., Aula N., Baerens C.E.M.T., Havelaar A.C., Voosge M., Peltonen L., Aula P., Galjaard H., Van der Spek P.J., Mancini G.M.S.;				
RT	"A new gene, encoding an anion transporter, is mutated in sialic acid storage diseases.";				
RL	Nat. Genet. 23:462-465(1999).				
RN	[3]				
RP	NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 2).				
RX	PubMed=14702039; DOI=10.1038/ng1285;				
RA	Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R., Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H., Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S., Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahari K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M., Tanai H., Kimura M., Watanabe M., Hirooka S., Chiba Y., Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Horita T., Kusunoki K., Takahashi F., Fujii A., Hara H., Tanase T.-O., Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N., Miasashiro K., Yuuki H., Oshima A., Sasaki N., Aotsuka S., Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,				

RA Moriya S., Momiyama H., Satoh N., Takemi S., Terahashi Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kamagaki A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki H., Hirao M., Ohmori Y.,
RA Kawabata A., Hiki T., Kobayashi N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Saeki M.,
RA Togaishi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuno Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isegaki T., Sugano S.,
RT "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs.",
RL Natl. Genet. 36:40-45(2004).
RN [4]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 2).
RP TISSUE=Colon;
RP MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RC Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RX Klusner R.D., Collins F.S., Wagner L., Shennen CM., Bhat N.K.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Burt R.K.,
RA Hopkins R.F., Jordan H., Moore T., Max I., Wang J., Hsieh F.,
RA DiCicco L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stadlton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rana S.S., Loguelfano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gamaralle P.H.,
RA Richards S., Wooley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kelleman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Rodriguez J.S.N., Krzyzanski M.I., Skalska U., Smalins D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra W.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.",
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [5]
RN VARIANTS SD CYS-39 AND GLU-136, AND VARIANTS ISSD 268-SER-ASN-272
RP DEL; ARG-183; ARG-334 AND VAL-371.
RX PubMed=10947946;
RA Aulia N., Salomaeaki P., Timonen R., Verheijen F., Mancini G.M.S.,
RA Maensson J.-E., Aulia P., Pelltonen L.,
RT "The spectrum of SLC17A5-gene mutations resulting in free static acid-
RT storage diseases indicates some genotype-phenotype correlation.",
RL Am. J. Hum. Genet. 67:832-840(2000).
RN [6]
RN VARIANT SD CYS-39.
RX MEDLINE=22678449; PubMed=12794687; DOI=10.1002/ajmg.a.10246;
RA Martin R.A., Slaugh R., Natowicz M., Pearlman K., Orvisky E.,
RA Krasnewich D., Kleiter R., Huizinga M., Gahl W.A.;
RT "Sialic acid storage disease of the Salla phenotype in American
RT monozygous twin female sibs.",
RL Am. J. Med. Genet. A 120:33-37(2003).
CC -1- FUNCTION: Primary solute translocator for anionic substances;
CC particularly it is a free static acid transporter in the lysosomes
CC (Probable).
CC -1- SUBCELLULAR LOCATION: Lysosome; Lysosomal membrane; multi-pass
CC membrane protein (Probable).
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=G9NRA2-1; Sequence=Displayed;
CC Name=2;
CC IsoId=G9NRA2-2; Sequence=VSP_010482, VSP_010483;
CC Note=NO experimental confirmation available;
CC -1- TISSUE SPECIFICITY: Found in fetal lung and small intestine, and
CC at lower level in fetal skin and muscle. In the adult, detected in
CC placenta, kidney and pancreas. Abundant in the endothelial cells
CC of tumors from ovary, colon,breast and lung, but is not detected

[illegible]

FT VARIANT 371 371 G -> V (in ISSD).
/FTID=VAR_018689.

Query Match 92.4%; Score 2621; DB 1; Length 495;
Best Local Similarity 100.0%; Pred. No. 1.9e-190;
Matches 495; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 42 MRSFVRDLARNDEESTDRTPLLPGAPRAEAAPVCCSARVNLALIAFFGFPIVVALRVNL 101
Db 1 MRSFVRDLARNDEESTDRTPLLPGAPRAEAAPVCCSARVNLALIAFFGFPIVVALRVNL 60

Qy 102 SVALVDMVDSNTTLEDNRTSKACEHSAPIKVHHNQGKKYQMDAETQWILGSPFYGYI 161
Db 61 SVALVDMVDSNTTLEDNRTSKACEHSAPIKVHHNQGKKYQMDAETQWILGSPFYGYI 120

Qy 162 ITQIPGGYVASKIGGKMLGFGILGTAVLTLFTPIADLGVPLIVRALGEGVTFP 221
Db 121 ITQIPGGYVASKIGGKMLGFGILGTAVLTLFTPIADLGVPLIVRALGEGVTFP 180

Qy 222 AMHAMSSMAPPLERSKLTLSYAGAQLGTVISLPLSGIICYMMNTYVFYFPGTIGIFW 281
Db 181 AMHAMSSMAPPLERSKLTLSYAGAQLGTVISLPLSGIICYMMNTYVFYFPGTIGIFW 240

Qy 282 FLMIWLVSDTPQGHKRI SHYEKEYIISLRNQLSSOKSVWPVILKSLPLMAIVAHFS 341
Db 241 FLMIWLVSDTPQGHKRI SHYEKEYIISLRNQLSSOKSVWPVILKSLPLMAIVAHFS 300

Qy 342 YNNTFYTLTLPLPYMKEILRFNVQENGFLSSLPYLSGMLCMILSGQAADNLRKKNFST 401
Db 301 YNNTFYTLTLPLPYMKEILRFNVQENGFLSSLPYLSGMLCMILSGQAADNLRKKNFST 360

Qy 402 LCVARRISLIGMIGPAFLVAAGFIGDYSLAVALFTISTTLGGFCSSGFSINHLDIAPS 461
Db 361 LCVARRISLIGMIGPAFLVAAGFIGDYSLAVALFTISTTLGGFCSSGFSINHLDIAPS 420

Qy 462 YAGILGITNTFATIPGMVGPVIAKSLTPDNTVGEMQTVFYIAAIVFGAIFFTLPAKG 521
Db 421 YAGILGITNTFATIPGMVGPVIAKSLTPDNTVGEMQTVFYIAAIVFGAIFFTLPAKG 480

Qy 522 EVQNMALNDHGHGRH 536
Db 481 EVQNMALNDHGHGRH 495

RESULT 2

Q55276_HUMAN PRELIMINARY; PRT; 495 AA.
AC Q55276;
DT 10-MAY-2005, integrated into UniProtKB/TrEMBL.
DT 10-MAY-2005, sequence version 1.
DE 07-FEB-2006, entry version 9.
DE Solute carrier family 17 (Anion/\sugar transporter), member 5.
GN Name=SLC17A5; ORFNames=RP3-397H23.3-001;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominiidae;
OC Homo
NCBI_TaxID=9606;
RN NCB1
RX NCB1_TaxID=9606;
RA NUCLEOTIDE SEQUENCE.
RA Chapman J.;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
RN NUCLEOTIDE SEQUENCE.
RA Kay M.;
RL Submitted (MAY-2005) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonDerivative License
CC EMBL; AL121972; CA120417.1; -; Genomic DNA.
CC EMBL; AL121972; CA120417.1; JOINED; Genomic DNA.
CC EMBL; AL121972; CA120417.1; JOINED; Genomic DNA.
DR EMBL; AL121972; CA120417.1; -; Genomic DNA.
DR EMBL; AL121972; CA120417.1; JOINED; Genomic DNA.
DR EMBL; AL121972; CA120417.1; JOINED; Genomic DNA.

DR EMBL; AL590428; CA115635.1; -; Genomic DNA.
DR Ensemble; ENSG00000119899; Homo sapiens.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0016020; C: membrane; IEA.
DR GO; GO:0005351; F: sugar porter activity; IEA.
DR GO; GO:0005215; F: transporter activity; IEA.
DR GO; GO:0006810; P: transport; IEA.
DR InterPro; IPR007114; MPS.
DR InterPro; IPR011701; MPS_1.
DR Pfam; PF07690; MPS_1.
DR PROSITE; PS50850; MPS; 1.
DR Membrane; Sugar transporter; Transmembrane; Transport.
SQ SEQUENCE 495 AA; 54640 MW; 5C6C154B3E93A19E CRC64;

Query Match 92.4%; Score 2621; DB 2; Length 495;
Best Local Similarity 100.0%; Pred. No. 1.9e-190;
Matches 495; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 42 MRSFVRDLARNDEESTDRTPLLPGAPRAEAAPVCCSARVNLALIAFFGFPIVVALRVNL 101
Db 1 MRSFVRDLARNDEESTDRTPLLPGAPRAEAAPVCCSARVNLALIAFFGFPIVVALRVNL 60

Qy 102 SVALVDMVDSNTTLEDNRTSKACEHSAPIKVHHNQGKKYQMDAETQWILGSPFYGYI 161
Db 61 SVALVDMVDSNTTLEDNRTSKACEHSAPIKVHHNQGKKYQMDAETQWILGSPFYGYI 120

Qy 162 ITQIPGGYVASKIGGKMLGFGILGTAVLTLFTPIADLGVPLIVRALGEGVTFP 221
Db 121 ITQIPGGYVASKIGGKMLGFGILGTAVLTLFTPIADLGVPLIVRALGEGVTFP 180

Qy 222 AMHAMSSMAPPLERSKLTLSYAGAQLGTVISLPLSGIICYMMNTYVFYFPGTIGIFW 281
Db 181 AMHAMSSMAPPLERSKLTLSYAGAQLGTVISLPLSGIICYMMNTYVFYFPGTIGIFW 240

Qy 282 FLMIWLVSDTPQGHKRI SHYEKEYIISLRNQLSSOKSVWPVILKSLPLMAIVAHFS 341
Db 241 FLMIWLVSDTPQGHKRI SHYEKEYIISLRNQLSSOKSVWPVILKSLPLMAIVAHFS 300

Qy 342 YNNTFYTLTLPLPYMKEILRFNVQENGFLSSLPYLSGMLCMILSGQAADNLRKKNFST 401
Db 301 YNNTFYTLTLPLPYMKEILRFNVQENGFLSSLPYLSGMLCMILSGQAADNLRKKNFST 360

Qy 402 LCVARRISLIGMIGPAFLVAAGFIGDYSLAVALFTISTTLGGFCSSGFSINHLDIAPS 461
Db 361 LCVARRISLIGMIGPAFLVAAGFIGDYSLAVALFTISTTLGGFCSSGFSINHLDIAPS 420

Qy 462 YAGILGITNTFATIPGMVGPVIAKSLTPDNTVGEMQTVFYIAAIVFGAIFFTLPAKG 521
Db 421 YAGILGITNTFATIPGMVGPVIAKSLTPDNTVGEMQTVFYIAAIVFGAIFFTLPAKG 480

Qy 522 EVQNMALNDHGHGRH 536
Db 481 EVQNMALNDHGHGRH 495

RESULT 3

S17A5 SHEEP STANDARD; PRT; 495 AA.
ID S17A5_SHEEP
AC Q9WZD1;
DT 07-JUN-2004, integrated into UniProtKB/Swiss-Prot.
DT 01-OCT-2000, sequence version 1.
DT 07-FEB-2006, entry version 22.
DE Sialin (solute carrier family 17 member 5) (sodium/sialic acid cotransporter) (Membrane glycoprotein SP55).
GN Name=SLC17A5;
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
OC Pecora; Bovidae; Caprinae; Ovis.
NCBI_TaxID=9940;
RN NCB1
RX NCB1_TaxID=9940;
RA NUCLEOTIDE SEQUENCE [MRNA], FUNCTION, AND TISSUE SPECIFICITY.
RP TISSUE=Lung;
CC

RX PubMed:11751519:
 Fu C., Bardhan S., Cateanu N.D., Wami B.D., Wang Y., Yan H.-P.,
 RA Shi E., Carter C., Venkov C., Yakes F.M., Page D.L., Lloyd R.S.,
 RA Merriam R.L., Hellerqvist C.G.,
 RT Identification of a novel membrane protein, HP59, with therapeutic
 potential as a target of tumor angiogenesis.";
 RL Clin. Cancer Res. 7:4182-4194(2001).
 CC -1- FUNCTION: Primary solute translocator for anionic substances;
 CC particularly it is a free fatty acid transporter in the lysosomes
 CC (probable). Receptor for CM101, a polysaccharide produced by group
 CC B Streptococcus with antipathogenic properties.
 CC -1- SUBCELLULAR LOCATION: Lysosome; lysosomal membrane; multi-pass
 CC membrane protein (potential).
 CC -1- TISSUE SPECIFICITY: Significantly expressed in lung endothelial
 CC cells, and much less in liver.
 CC -1- SIMILARITY: Belongs to the major facilitator superfamily.
 CC Sodium/anion cotransporter family.
 CC -----
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonDerivative license
 CC -----
 CC EMBL: AF244578; AAF97770.1; -; mRNA.
 DR InterPro: IPR007114; MFS.
 DR InterPro: IPR011701; MFS_1.
 DR Pfam: PF07690; MFS_1; 1.
 DR PROSITE: PS0850; MFS_1.
 KW Lysosome; Membrane; Receptor; Symport; Transmembrane; Transport.
 KM
 FT CHAIN 1 495
 FT TRANSMEM 42 62 /FTid=PRO_0000220949.
 FT TRANSMEM 110 130 Potential.
 FT TRANSMEM 137 157 Potential.
 FT TRANSMEM 159 179 Potential.
 FT TRANSMEM 201 221 Potential.
 FT TRANSMEM 228 248 Potential.
 FT TRANSMEM 289 309 Potential.
 FT TRANSMEM 329 349 Potential.
 FT TRANSMEM 366 386 Potential.
 FT TRANSMEM 392 412 Potential.
 FT TRANSMEM 424 444 Potential.
 FT TRANSMEM 458 478 Potential.
 SQ SEQUENCE 495 AA; 54536 MW; 649D7CA59B28272 CRC64;
 Query Match 82.1%; Score 2329; DB 1; Length 495;
 Best Local Similarity 86.3%; Pred. NO. 2.9e-168;
 Matches 427; Conservative 39; Mismatches 29; Indels 0; Gaps 0;
 42 MRSPPVDLANDEESTDRTPLPGAPRAEAPVCCSARYNLAIIAFEGFIVYALRVNL 101
 1 MKSPVDLAPSDDESGDRTPLQGRAPRAPVCCSARYNLAFLSFGFVLYSLRVNL 60
 102 SVALVMDVDSNTLLEDRKTSKACPEHSAPIKVAHNOGKYKQMDAQOMILSFFGYI 161
 61 SVALVMDVDSNTAKDNRTSYECASHAPIKVLNNOGKYKQMDAQOMILSFFGYI 120
 162 ITQIPGGYVASKGKMLGFGILGTAVLTFPIADLGVPLIVRALEGGEVTP 221
 121 ITQIPGGYVSRSGKLLGFGIPATVIFLTFPLADFGVALVALRALEGGEVTP 180
 222 AMHAMSSMAPLERSKLISVAGAGLVISLPLSGIICYYNMNTYVFFGTTIGIFW 281
 181 AMHAMSSMAPLERSKLISVAGAGLVISLPLSGIICYYNMNTYVFFGTTIGIFW 240
 282 FLIMLVSTDPQGHKRIKSHYKVIISLRNQLSSQKSPVWPVITLKLPLMAIVAHFS 341
 241 FLIMLVSTDPQGHKRIKSHYKVIISLRNQLSSQKSPVWPVITLKLPLMAIVAHFS 300
 342 YNNTFYTLTLPLTYMEIIRFNVQENGLSPYLGLSWICMISGQAAADLRAKNMFT 401
 301 YNNTFYTLTLPLTYMEIIRFNVQENGLSPYLGLSWICMISGQAAADLRAKNMFT 360
 402 LCVRIFFSLIGMIGPAVFLVAAAGFICGDYSLAVAFLTISTTLGGFCSSGSINHLDIAPS 461

Db 361 LWRVRVFSLIGMIGPAVFLVAAAGFICGDYSLAVAFLTISTTLGGFCSSGSINHLDIAPS 420
 Qy 462 YGIIILGINTPATTPGMWGVIAKSLTPDNTVGEMQTVFYAAAINVGAIFFTLFANG 521
 Db 421 YGIIILGINTPATTPGMWGVIAKSLTPDNTVGEMQTVFYAAAINVGAIFFTLFANG 480
 Qy 522 EVONMALNDHGHHRH 536
 Db 481 EVONMALNDHGHHRH 495
 RESULT 4
 S17A5 MOUSE
 ID S17A5 MOUSE STANDARD; PRT; 495 AA.
 AC OBBN82;
 DT 07-JUN-2004, integrated into UniProtKB/Swiss-Prot.
 DT 07-JUN-2004, sequence version 21.
 DE 07-FEB-2006, entry version 21.
 DE Stalim (Solute carrier family 17 member 5) (Sodium/stalic acid
 DE cotransporter).
 GN Name=Slc17a5;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognath;
 OC Muridae; Muridae; Murinae; Mus.
 OX NCBI_TaxId=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORMS 1 AND 3).
 RC STRAIN=CS7BL/6J; TISSUE=Eye, and Skin.
 RX PubMed=16141072; DOI=10.1126/science.1112014.
 RA Carninci P., Kasukawa T., Katayama S., Gough J., Fritch M.C., Maeda N.,
 RA Oyama R., Ravasi T., Lenhard B., Wells C., Kodzius R., Shimokawa K.,
 RA Bajic V.B., Brenner S.E., Batalov S., Forrest A.R., Zavolan M.,
 RA Davis M.J., Wilm M., Delventhal L.G., Altieri F., Allen J.E.,
 RA Ambesi-Impombato A., Apweiler R., Aturaliya R.N., Bailey T.L.,
 RA Bansal M., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M.,
 RA Chiu K.P., Choudhary V., Christoffels A., Clutterbuck D.R.,
 RA Crowe M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,
 RA di Bernardo D., Down T., Engstrom P., Fagioli M., Faulkner G.,
 RA Fletcher C.F., Fukushima T., Furuno M., Futaki S., Gariboldi M.,
 RA Georgii-Hemming P., Gingeras T.R., Gojobori T., Green R.E.,
 RA Guerginich S., Hachisu T., Hayes A., Hensch T.K., Hirokawa N.,
 RA Hill D., Hummelich A., Iacono M., Ikeo K., Iwama A., Ishikawa T.,
 RA Jaki M., Kanapin A., Katoh M., Kawasawa Y., Kello J., Kitamura H.,
 RA Kizuno H., Kohara A., Krishnan S.P., Kruger A., Kummerfeld S.K.,
 RA Kurochkin I.V., Lareau L.F., Lazarevic D., Lipovich L., Liu J.,
 RA Liuni S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,
 RA Matsuda M., Matsuzawa S., Miki H., Mignone F., Miyake S., Morris K.,
 RA Mottagui-Tabar S., Mulder N., Nakano N., Nakachi H., Ng P.,
 RA Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,
 RA Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavoni G., Pesole G.,
 RA Petrovsky N., Piazza S., Reed J., Reid J.F., Ring B.Z., Ringwald M.,
 RA Rostrovsky N., Salzberg S.L., Sanderlin A., Schneider C.,
 RA Schombach C., Sekiguchi K., Sempile C.A., Seno S., Sessa L., Sheng Y.,
 RA Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,
 RA Sperling S., Stupka E., Sugita K., Sultana R., Takenaka Y., Taki K.,
 RA Tamura K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,
 RA Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yang J.,
 RA Yamashiki H., Zdobych E., Zhu S., Zimmer A., Hilde W., Bult C.,
 RA Grimmond S.M., Teasdale R.D., Liu E.T., Brusic V., Quekembush J.,
 RA Wahlestedt C., Mattick J.S., Hume D.A., Kai C., Saeki K., Tomaru Y.,
 RA Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Aizawa T.,
 RA Iida J., Imamura K., Itoh K., Kato T., Kawasaki H., Kawagashira N.,
 RA Kawashima T., Kojima M., Kondo S., Konno H., Nakano K., Niimura N.,
 RA Niehio T., Okada M., Pleasy C., Shibata K., Shireki T., Suzuki S.,
 RA Tegen M., Waki K., Watanuki A., Okamura-Oho Y., Suzuki H., Kawai J.,
 RA Hayashizaki Y.,
 RT "The transcriptional landscape of the mammalian genome.";
 RL Science 309:1559-1563(2005).
 RN [2]
 RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] (ISOFORM 2).
 RC TISSUE=Olfactory epithelium;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;

RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Satchell M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udell T.B., Toshiyuki S., Carrincci P., Prange C.,
 RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mulhally S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Wotley K.C., Hale S., Garcia A.M., Gay L.J., Halyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fleury J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Buttefield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
 RA Schenck A., Schein J.B., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- FUNCTION: Primary soluble translocator for anionic substances;
 CC particularly it is a free sialic acid transporter in the lysosomes
 CC (probable).
 CC -1- SUBCELLULAR LOCATION: Lysosome; lysosomal membrane; multi-pass
 CC membrane protein (Potential).
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=3;
 CC Name=1;
 CC IsoId=Q8BN82-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=Q8BN82-2; Sequence=VSP_010484;
 CC Note=No experimental confirmation available;
 CC Name=3;
 CC IsoId=Q8BN82-3; Sequence=VSP_010485, VSP_010486;
 CC Note=No experimental confirmation available;
 CC -1- SIMILARITY: Belongs to the major facilitator superfamily.
 CC Sodium/anion cotransporter family.
 CC -----
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonDerivs License
 CC -----
 CC EMBL: AK029102; BAC26298.1; -; mRNA.
 CC EMBL: AK087395; BAC39859.1; -; mRNA.
 CC EMBL: BC058785; AAH58785.1; -; mRNA.
 CC Ensembl: ENSMUSG00000049624; Mus musculus.
 CC MGI: MGI:1924105; Slc17a5.
 CC GO: GO:0016023; C:cytoplasmic membrane-bound vesicle; IDA.
 CC GO: GO:0005886; C:plasma membrane; IDA.
 CC InterPro: IPR007114; MFS.
 CC InterPro: IPR011701; MFS_1.
 CC Pfam: PF07690; MFS_1; 1.
 CC PROSITE: PSS0850; MFS; 1.
 CC Alternative splicing; Lysosome; Membrane; Symport; Transmembrane;
 CC Transport.
 CC CHAIN 1 495 Sialin.
 FT TRANSMEM 42 62 /Frtid=PRO_0000220948.
 FT TRANSMEM 110 130 Potential.
 FT TRANSMEM 137 157 Potential.
 FT TRANSMEM 159 179 Potential.
 FT TRANSMEM 201 221 Potential.
 FT TRANSMEM 228 248 Potential.
 FT TRANSMEM 279 299 Potential.
 FT TRANSMEM 329 349 Potential.
 FT TRANSMEM 366 386 Potential.
 FT TRANSMEM 392 412 Potential.
 FT TRANSMEM 424 444 Potential.
 FT TRANSMEM 458 478 Potential.
 FT VARSPPLIC 98 123 Missing (in isoform 2).
 FT VARSPPLIC 176 238 /Frtid=VSP_010484
 FT GVTPEAHAMSSWAPPLERSKLTITSYAGQLGTVISLPL
 FT SGIICYYMMNTYVYFGLIGVI -> KYPFGCVSSYARH
 FT VVPLGSPGKKQASVHFLCGSTAWDSLTSSFRNNMLYEL

FT FT DLRLSP (in isoform 3).
 FT FT /Frtid=VSP_010485.
 FT VARSPPLIC 239 495 Missing (in isoform 3).
 FT SO SEQUENCE 495 AA; 54369 MW; 6009661215D26437 CRC64;
 Query Match 81.0%; Score 2296; DB 1; Length 495;
 Best Local Similarity 86.3%; Pred. No. 6.5e-16;
 Matches 427; Conservative 32; Mismatches 36; Indels 0; Gaps 0;
 QY 42 MRSFVRDLANNDEESTDRTPLLPGAPRAAPYCCSARYNALALAFPGFFIYALRVNL 101
 DB 1 MRPLLRPGAGNDEESSDSTPLPGARQTEAAPCCSARYNALALAFCGFFVLALRVNL 60
 QY 102 SVLAIVDVNDNTLEDNRTKACGEHAPLKVNHNQKKYQMDAEFGMTLGSFFGYI 161
 DB 61 SVALVDVNDNTLTDRTSKCEHSAPIKVNHNQKKYKMDAEFGMTLGSFFGYI 120
 QY 162 ITOIPGVYVASKIGKMLGFGILGTAVLTLPPIADLGVPPIVRLALEGLGEVTFP 221
 DB 121 VTQIPGVYVASKIGKMLGFGILGTAVLTLPPIADLGVPPIVRLALEGLGEVTFP 180
 QY 222 AMHAMSSWAPPLERSKLTITSYAGQLGTVISLPLSGIICYYMMNTYVYFGLIGVI 281
 DB 181 AMHAMSSWAPPLERSKLTITSYAGQLGTVISLPLSGIICYYMMNTYVYFGLIGVI 240
 QY 282 FLIMWIVSDTPQKHRIHYEKEXYLSLRNQLSSQKSPWVPIILKSLPLMAIVVAHFS 341
 DB 241 FLIMWIVSDTPQKHRIHYEKEXYLSLRNQLSSQKSPWVPIILKSLPLMAIVVAHFS 300
 QY 342 YNNTFYTLTLTLPTYMKELIRFNVQENGFLSYLPSYLSWMLSCMLSGAADNLRKWNFT 401
 DB 301 YNNTFYTLTLTLPTYMKELIRFNVQENGFLSYLPSYLSWMLSCMLSGAADNLRKWNFT 360
 QY 402 LCVRRIFSLGIMGPAVLVAAGFTGCDYSLAVFLTISTTLGFCSSGFSINHLDIAPS 461
 DB 361 ISVRRIFSLGIMGPAVLVAAGFTGCDYSLAVFLTISTTLGFCSSGFSINHLDIAPS 420
 QY 462 YAGILGINTPTPTIRPMGVPVIAKSLTPNTYNGEMQTVYIAAIVFGAIFPTLRPKG 521
 DB 421 YAGILGINTPTPTIRPMGVPVIAKSLTPNTYNGEMQTVYIAAIVFGAIFPTLRPKG 480
 QY 522 EVQNMALNDHGRHRH 536
 DB 481 EVQNMALNDHGRHRH 495
 RESULT 5
 Q3TE25_MOUSE PRELIMINARY; PRT; 495 AA.
 ID Q3TE25_MOUSE Integrated into UniProtKB/TrEMBL.
 AC 11-OCT-2005, integrated into UniProtKB/TrEMBL.
 DT 11-OCT-2005, sequence version 1.
 DT 07-FEB-2006, entry version 7.
 DE NOD-derived CD1c +ve dendritic cells cDNA, RIKEN full-length enriched
 DE library, clone: F630001P18 product: soluble carrier family 17
 DE (anion/sugar transporter), member 5, full insert sequence.
 GN Name=Slc17a5;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP NOCLOTIDE SEQUENCE.
 RC STRAIN=NOD;
 RX MEDLINE=9279253; Pubmed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
 RA Carninci P., Hayashizaki Y.;
 RT "High-efficiency full-length cDNA cloning.";
 RU Methods Enzymol. 303:19-44(1999).
 RN [2]
 RP NOCLOTIDE SEQUENCE.
 RC STRAIN=NOD;

RA PubMed:16141073; DOI=10.1126/science.1112014;
RA Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N.,
RA Oyama R., Ravasi T., Lenhard B., Wells C., Kodzius R., Shimokawa K.,
RA Bajic V.B., Bremner S.E., Batalov S., Forrest A.R., Zavolan M.,
RA Davis M.J., Wilming L.G., Aldins J., Allen J.E.,
RA Ambeyi-Imbombo A., Appleier R., Aturaliya R.N., Bailey T.L.,
RA Banasi M., Baxter L., Beisel K.W., Berano T., Bono H., Chalk A.M.,
RA Chiu K.P., Choudhary V., Christoffels A., Cluterebuck D.R.,
RA Crowe M.L., Dalla E., Dalrymple B.P., de Bono B., Della Gatta G.,
RA di Bernardo D., Down T., Engstrom P., Fagioli M., Faulkner G.,
RA Fletcher C.F., Fukushima T., Furuno M., Futaki S., Gariboldi M.,
RA Georgil-Heming P., Gingeras T.R., Gojobori T., Green R.E.,
RA Gustinich S., Harbers M., Hayashi Y., Hensch T.K., Hirokawa N.,
RA Hill D., Hummel L., Iacono M., Ikeo K., Iwama A., Ishikawa T.,
RA Jakt M., Kanapin A., Katoh M., Kawasawa Y., Kelso J., Kikunaga H.,
RA Kikano H., Kollias G., Krishnan S.P., Kruger A., Kumfeld S.K.,
RA Kurochkin I.V., Lareau L.F., Lazarevic D., Lipovich L., Liu J.,
RA Liuni S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,
RA Matsuda H., Matsuzawa S., Miki H., Mignone F., Miyake S., Morris K.,
RA Mottegu-Taber S., Mulder N., Nakano N., Nakachi H., Ng P.,
RA Nilsen R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,
RA Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavoni G., Pesole G.,
RA Petrovsky N., Piazza S., Reed J., Reid J.F., Ring B.Z., Ringwald M.,
RA Rost B., Ruan Y., Salzberg S.L., Sanderlin A., Schneider C.,
RA Schombach C., Sekiguchi K., Sempole C.A., Seno S., Sees L., Sheng Y.,
RA Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,
RA Sperling S., Stupka E., Sugita K., Sultana R., Takenaka Y., Taki K.,
RA Tammoja K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,
RA Ueda H.R., van Nimwegen E., Verardo R., Wei C.L., Yeai K.,
RA Yamashita H., Zabarovsky E., Zhu S., Zimmer A., Hide W., Bult C.,
RA Grimmond S.M., Teasdale R.D., Liu E.T., Brusic V., Quackenbush J.,
RA Wahlstedt C., Matic J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,
RA Fukuda S., Kanamori-Katayama M., Suzuki M., Aoki J., Arakawa T.,
RA Iida J., Imamura K., Itoh M., Kato T., Kawaji H., Kawaguchi N.,
RA Kawashima T., Kojima M., Kondo S., Konno H., Nakano K., Nimomiya N.,
RA Nishio T., Okada M., Plessey C., Shibata K., Shiraki T., Suzuki S.,
RA Tagami M., Waki K., Watabiki A., Okamura-Oho Y., Suzuki H., Kawai J.,
RA Hayashizaki Y.,
RT "The transcriptional landscape of the mammalian genome.",
RL Science 309:1559-1563(2005).
[3]
RP NUCLEOTIDE SEQUENCE.
RP STRAIN=ND;
RC PubMed:16141073; DOI=10.1126/science.1112009;
RG RIKEN Genome Exploration Research Group and Genome Science Group
RT (Genome Network Core Team) and the FANTOM Consortium.
RL "Antisense transcription in the mammalian transcriptome.",
RL Science 309:1564-1566(2005).
[4]
RP NUCLEOTIDE SEQUENCE.
RP STRAIN=ND;
RC MEDLINE=22354683; PubMed=1246685; DOI=10.1038/nature01266;
RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
RA Nakajima I., Osada N., Saito R., Suzuki H., Yamashita H., Kiyosawa H.,
RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schombach C., Gojobori T.,
RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
RA Blake J.A., Bratt D., Bruce V., Chochia C., Corbett L.E., Cousins S.,
RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,
RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
RA Grimmond S., Gustinich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
RA Kanai A., Kawai H., Kawasawa Y., Kedzierski R.M., King B.L.,
RA Konegaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
RA Meglert D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
RA Nagashima T., Numata K., Okido T., Pavan W.J., Petrea G., Pesole G.,
RA Petrovsky N., Pillai R., Pontius J.V., Qi D., Ramachandran S.,
RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
RA Sanderlin A., Schneider C., Sempole C.A., Setou M., Shimada K.,
RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
RA Verardo R., Wagner L., Wahlstedt C., Wang Y., Watanabe Y., Wells C.,
RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,

RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Iishi Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shingawa A.,
RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Birney E., Hayashizaki Y.,
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.",
RL Nature 420:563-573(2002).
[5]
RP NUCLEOTIDE SEQUENCE.
RP STRAIN=ND;
RC MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Iishi Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamashita H.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nishio T., Pesole G., Quackenbush J.,
RA Schriml L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barash G.,
RA Blake J., Beffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustinich S., Hill D., Holman M., Hume D.A., Kamaya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaele P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schombach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohseki S.,
RA Hayashizaki Y.,
RT "Functional annotation of a full-length mouse cDNA collection.",
RL Nature 409:685-690(2001).
[6]
RP NUCLEOTIDE SEQUENCE.
RP STRAIN=ND;
RC MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.,
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.",
RL Genome Res. 10:1617-1630(2000).
[7]
RP NUCLEOTIDE SEQUENCE.
RP STRAIN=ND;
RC PubMed:20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama U., Nishi K., Kikunaga T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watabiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsumura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.,
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.",
RL Genome Res. 10:1757-1771(2000).
[8]
RP NUCLEOTIDE SEQUENCE.
RP STRAIN=ND;
RC Arakawa T., Carninci P., Fukuda S., Hashizume W., Hayashida K.,
RA Hori F., Iida J., Imamura K., Imotani K., Itoh M., Kanagawa S.,
RA Kawai J., Kojima M., Konno H., Murata M., Nakamura M., Nimomiya N.,
RA Nishiyori H., Nomura K., Ohno M., Sakakume Y., Sano H., Sasaki D.,
RA Shibata K., Shiraki T., Tagami M., Tagami Y., Waki K., Watabiki A.,
RA Muramatsu M., Hayashizaki Y.,
RT Submitted (Apr-2004) to the EMBL/GenBank/DBJ databases.
RL
CC Copyrighted by the Uniprot Consortium, see <http://www.uniprot.org/terms>
CC Distributed under the Creative Commons Attribution-NonDerivative license
CC
CC EMBL: AK169688; BAE41423.1; -; mRNA.
MG: MGI:1924105; SLC17A5.
DR GO: GO:0016023; C:cytoplasmic membrane-bound vesicle; IDA.
GO: GO:0005886; C:plasma membrane; IDA.

DR GO; GO:0005215; F:transporter activity; RCA.
 DR GO; GO:0006810; P:transport; RCA.
 DR InterPro: IPR007114; MFS.
 DR InterPro: IPR011701; MFS_1.
 DR Pfam: PF07690; MFS_1; 1.
 DR PROSITE: PS00850; MFS; 1.
 KM Membrane; Sugar transport; Transmembrane; Transport.
 SQ SEQUENCE 495 AA; 54369 MW; 6009661215D26437 CRC64;

Query Match 81.0%; Score 2298; DB 2; Length 495;
 Best Local Similarity 86.3%; Pred. No. 6.5e-166;
 Matches 427; Conservative 32; Mismatches 36; Indels 0; Gaps 0;

QY 42 MRSPVRLDARDGEESTDRTPLLPGAPRAEAPCCSARVNLALIAFGFFIYALRVNL 101
 1 MRPLRLRPGAGNDESSDSTPLPGARQTEAPPCCSARVNLALIAFGFFIYALRVNL 60
 DB SVALVDMVDSNTLTLDNRSTKCAPEHSAPIKVHNHNGKKYQMDAETQGMILGSFFGYI 120
 QY 102 SVALVDMVDSNTLTLDNRSTKCAPEHSAPIKVHNHNGKKYQMDAETQGMILGSFFGYI 161
 61 SVALVDMVDSNTLTLDNRSTKCAPEHSAPIKVHNHNGKKYQMDAETQGMILGSFFGYI 120
 DB 162 ITQIPGGYVASKIGGKMLGFGILGTRAVTLFPIAADLGVPILVIRALEGEGVTFP 221
 121 VTQIPGGYVASKIGGKMLGFGILGTRAVTLFPIAADLGVPILVIRALEGEGVTFP 180
 QY 222 AMHAMSSMAPPLERSKLISYAGAQGVISLPLSGIICYYNMNTYVFFGTIGIFW 281
 181 AMHAMSSMAPPLERSKLISYAGAQGVISLPLSGIICYYNMNTYVFFGTIGIFW 240
 DB 282 FLLMIWVSDTPQKHKRISHYEKEYILSLRNQSSQKVPWVPIILKSLPLMAIVAHS 341
 241 FILMMIVSDTPETHKTIISHYEKEYIVSLKNQSSQKVPWVPIILKSLPLMAIVAHS 300
 QY 342 YNNFTYTLTLPTPKKEILRFNVQENGFLSLPYLGSWLCMLISGQADNLRKKNFST 401
 301 YNNFTYTLTLPTPKKEILRFNVQENGFLSLPYLGSWLCMLISGQADNLRKKNFST 360
 DB 402 LCVRRIFSLIGMIGPAVFLVAGTIGCDYSIAVAFLLISTTLGGFCSSGFSINLADIAPS 461
 361 ISVRRISLVGMGPVFLVAGTIGCDYSIAVAFLLISTTLGGFCSSGFSINLADIAPS 420
 QY 462 YAGILGITNTFATIPGMVGPVIAKSLTPDNTVGEQMTVFYIAAIVFCAIFPTLPAGK 521
 421 YAGILGITNTFATIPGMVGPVIAKSLTPDNTVGEQMTVFYIAAIVFCAIFPTLPAGK 480
 DB 522 EYQNMALNDHGHHRH 536
 481 EYQNMALNDHGHHRH 495

QY 522 EYQNMALNDHGHHRH 536
 481 EYQNMALNDHGHHRH 495

RESULT 6
 Q50000 RAT PRELIMINARY; PRT; 495 AA.
 AC Q50000;
 DT 04-JAN-2005, integrated into UniProtKB/TrEMBL.
 DT 04-JAN-2005, sequence version 1.
 DT 07-FEB-2006, entry version 13.
 DE Stailin (solute carrier family 17 (anion/sugar transporter), member 5).
 GN Name=Slc17a5;
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muroidae; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RN NUCLEOTIDE SEQUENCE.
 RP PubMed=15516337; DOI=10.1074/jbc.M411295200;
 RX Wreden C.C., Wlitzla M., Reimer R.J.;
 RT "Varied mechanisms underlie the free stailic Acid storage disorders";
 RJ J. Biol. Chem. 280:1408-1416(2005).
 RN [2]
 RN NUCLEOTIDE SEQUENCE.
 RP TISSUE=Placenta;

RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan J., Moore T., Max S.I., Wang J., Heien F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stajich M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smallie D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [3]
 RN NUCLEOTIDE SEQUENCE.
 RP TISSUE=Placenta;
 RG NIH MGC Project;
 RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonDerivs License

DR EMBL: AY800277; AA973775.1; -; mRNA.
 DR EMBL: BC097482; AA97482.1; -; mRNA.
 DR Ensembl: ENSRN00000009330; Rattus norvegicus.
 DR GO: GO:0016021; C:integral to membrane; IEA.
 DR GO: GO:0016020; C:membrane; IEA.
 DR GO: GO:0005351; F:sugar porter activity; IEA.
 DR GO: GO:0005215; F:transporter activity; IEA.
 DR GO: GO:0006810; P:transport; IEA.
 DR InterPro: IPR007114; MFS.
 DR InterPro: IPR011701; MFS_1.
 DR Pfam: PF07690; MFS_1; 1.
 DR PROSITE: PS00850; MFS; 1.
 KM Membrane; Sugar transport; Transmembrane; Transport.
 SQ SEQUENCE 495 AA; 54323 MW; AAB0A29BA5D42094 CRC64;

Query Match 81.0%; Score 2298; DB 2; Length 495;
 Best Local Similarity 86.8%; Pred. No. 6.5e-166;
 Matches 428; Conservative 29; Mismatches 36; Indels 0; Gaps 0;

QY 42 MRSPVRLDARDGEESTDRTPLLPGAPRAEAPCCSARVNLALIAFGFFIYALRVNL 101
 1 MRPLRLRPGAGNDESSDSTPLPGARQTEAPPCCSARVNLALIAFGFFIYALRVNL 60
 DB SVALVDMVDSNTLTLDNRSTKCAPEHSAPIKVHNHNGKKYQMDAETQGMILGSFFGYI 120
 QY 102 SVALVDMVDSNTLTLDNRSTKCAPEHSAPIKVHNHNGKKYQMDAETQGMILGSFFGYI 161
 61 SVALVDMVDSNTLTLDNRSTKCAPEHSAPIKVHNHNGKKYQMDAETQGMILGSFFGYI 120
 DB 162 ITQIPGGYVASKIGGKMLGFGILGTRAVTLFPIAADLGVPILVIRALEGEGVTFP 221
 121 VTQIPGGYVASKIGGKMLGFGILGTRAVTLFPIAADLGVPILVIRALEGEGVTFP 180
 QY 222 AMHAMSSMAPPLERSKLISYAGAQGVISLPLSGIICYYNMNTYVFFGTIGIFW 281
 181 AMHAMSSMAPPLERSKLISYAGAQGVISLPLSGIICYYNMNTYVFFGTIGIFW 240
 DB 282 FLLMIWVSDTPQKHKRISHYEKEYILSLRNQSSQKVPWVPIILKSLPLMAIVAHS 341
 241 FILMMIVSDTPETHKTIISHYEKEYIVSLKNQSSQKVPWVPIILKSLPLMAIVAHS 300
 QY 342 YNNFTYTLTLPTPKKEILRFNVQENGFLSLPYLGSWLCMLISGQADNLRKKNFST 401
 301 YNNFTYTLTLPTPKKEILRFNVQENGFLSLPYLGSWLCMLISGQADNLRKKNFST 360

```

QY 402 LCVRRIFSLIGMIGPAVFLVAAGFIGDYSILAVALFTISTTLGFCSSGFSINHLDIAPS 461
DB 361 ISRRIFSLIGMIGPAVFLVAAGFIGDYSILAVALFTISTTLGFCSSGFSINHLDIAPS 420
QY 462 YAGILGTTTFAITPGMVPVIAKSLTPDNTVGEMQTVFYIAAIVFGAIFFTLPAKG 521
DB 421 YAGILGTTTFAITPGMVPVIAKSLTPDNTIKEMQTVFYIAAIVFGAIFFTLPAKG 480
QY 522 EVONMALNDHGH 534
DB 481 EVONMALSDHGH 493

RESULT 7
Q5ZL94_CHICK PRELIMINARY; PRT; 484 AA.
AC 05ZL94;
DT 23-NOV-2004, integrated into UniProtKB/TrEMBL.
DT 23-NOV-2004, sequence version 1.
DT 21-FEB-2006, entry version 13.
DE Hypothetical protein.
GN ORFNames=RCJMB04_7b17;
OC Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CB; TISSUE=Bursa;
RA Caldwell R.B., Kierzek A.M., Arakawa H., Bezzubov Y., Zaim J.,
RA Fiedler P., Kuter S., Blagovska D., Koter M.,
RA Plachy J., Carnici P., Hayashizaki Y., Buerstedt J.M.,
RT "Full-length cDNAs from chicken Bursa lymphocytes to facilitate
RT gene function analysis."
RL Genome Biol. 6:R6-R6(2005).
CC -1- SUBCELLULAR LOCATION: Membrane; multi-pass membrane protein (By
CC similarity).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DB EMBL: AJ719840; CAG31499.1; -, mRNA.
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0005215; C:membrane; IEA.
DR GO: GO:0006810; P:transporter activity; IEA.
DR InterPro: IPR007114; MFS.
DR InterPro: IPR011701; MFS_1.
DR Pfam: PF07690; MFS_1; 1.
KW PROSITE: PS50850; MFS; 1.
KW Hypothetical protein; Membrane; Transmembrane; Transport.
SQ SEQUENCE 484 AA; 53557 MW; E9BDEE09FES2D502 CRC64;

Query Match 73.1%; Score 2072.5; DB 2; Length 484;
Best Local Similarity 79.5%; Pred. No. 8.6e-149;
Matches 383; Conservative 47; Mismatches 51; Indels 1; Gaps 1;

```

```

QY 295 KHKRI SHYEKEYILSSLRNQLSOKSVPMVPIKSLPLMAIVAHFSYMMPTTLTLTP 354
DB 243 THKSIIRIEREYILSSLRNQLSOKSVPMRPILESPLMAIVAHFSYMMPTTLTLTP 302
QY 355 TYMKELIRFNVQENGFLSSLPYGSWLCMILSGAANDIRAKNFTLCVRRIISLIGMI 414
DB 303 TYMKELIRFPAQENGFLSALPYGCWLCIILSGQIDHLREKQNFSTVCVRKCFTLIGMI 362
QY 415 GPAVFLVAAGFIGCDYSILAVALFTISTTLGFCSSGFSINHLDIAPSYAGILGINTPA 474
DB 363 GPAVFLVAAGFIGCDYSILAVALFTISTTLGFCSSGFSINHLDIAPSYAGILGINTPA 422
QY 475 TIRGMVGPVIAKSLTPDNTVGEMQTVFYIAAIVFGAIFFTLPAKGEVONMALNDHGH 534
DB 423 TIRGMVGPVIAKSLTPDNTVGEMQTVFYIAAIVFGAIFFTLPAKGEVONMALNDHGH 482
QY 535 RH 536
DB 483 RH 484

RESULT 8
Q4SGF2_TETNG PRELIMINARY; PRT; 490 AA.
AC Q4SGF2;
DT 19-JUL-2005, integrated into UniProtKB/TrEMBL.
DT 19-JUL-2005, sequence version 1.
DT 07-FEB-2006, entry version 4.
DE Chromosome 17 SCARF4597, whole genome shotgun sequence. (Fragment).
GN ORFNames=GSTEN00018669001;
OS Tetradon nigroviridis (Green puffer).
OC Tetradon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorphia; Acanthopterygii; Percormorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetradon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15496914; DOI=10.1038/nature03025;
RA Ujillon O., Aury J.-M., Brunet F., Petit J.-L., Stange-Thomann N.,
RA Maucell E., Bounneau L., Fischer C., Ozoul-Costaz C., Bernot A.,
RA Nicard S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segutene B.,
RA Desliiva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Jubin C., Castellio L., Poulain J., De Berardinis V.,
RA Biemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA Cruaud C., Duprat S., Brottier P., Coutanceau J.-P., Guzy J.,
RA Parra G., Lardier G., Chappe C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Volff J.-N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Landet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissbach J., Roest Crolius H.,
RT "Genome duplication in the teleost fish Tetradon nigroviridis reveals
RT the early vertebrate proto-karyotype."
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RG Genoscope, Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DB EMBL: CAE01014597; CAG00280.1; -, Genomic DNA.
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0005215; P:transporter activity; IEA.
DR GO: GO:0006810; P:transporter activity; IEA.
DR InterPro: IPR007114; MFS.
DR Pfam: PF07690; MFS_1; 1.
KW PROSITE: PS50850; MFS; 1.
DR

```

FT NON TER 1 1
 FT NON TER 490 490
 SQ SEQUENCE 490 AA; 54112 MM; C6724685C21A1B0D CRC64;

Query Match 55.8%; Score 1582.5; DB 2; Length 490;
 Best Local Similarity 59.5%; Pred. No. 1.5e-111;
 Matches 291; Conservative 74; Mismatches 97; Indels 27; Gaps 2;

QY 73 APVCSARVNLATIAFGFETVYALRVNLSVALVDMDSNTLTEDNRTSKACPEHSAPIK 132
 DB 1 APVCSARVNLATIAFGFETVYALRVNLSVALVDMDSNTLTEDNRTSKACPEHSAPIK 60
 QY 133 VHNQTKKRYQMDAETQGMILGSPFYGYITQIPGVYASKIGKMLGFGILGTAVLT 192
 DB 61 PRNQTASVYDMDSSETQGMILGAFYGYITQIPGVYASKIGKMLGFGILGTAVLT 120
 QY 193 FTPIADLVGVPILVLRALBGLGEGVTFPAMHAMSSWAPPLERSKLLSISYAGOLGTV 252
 DB 121 LTPILADLVGVPILVLRALBGLGEGVTFPAMHAMSSWAPPLERSKLLSISYAGOLGTV 180
 QY 253 ISLPLSGITCYMMWTVVFFFGITGIFWFLMTWLVSDTPQKAKRISHKEKYLISLR 312
 DB 181 IALPLSGITCYMMWTVVFFFGITGIFWFLMTWLVSDTPQKAKRISHKEKYLISLR 240
 QY 313 NQTS-SQKSVPMWPIKSLPLMAIVAHFSYNTFTLTLLPTYKEILRFNY----- 365
 DB 241 DELALSTHCLIPMAIVSRPLMAIVAHFPCNMSFTLLTLLPTYMDILGFSIQOQGN 300
 QY 366 -----QENGFLSSLPYLSGMLCMILSGQADNLRKWNFTSLCVR 405
 DB 301 ERKRPVMPAMTLKPRCLRPQNGMLBALPYLGCSIAVAFAGPADYLRCTLVSVR 360
 QY 406 RITSLSGIMIGPAFLVAGFGICDYSLAVAFITSTLTGSCSGSEINHLDIAPSYAGI 465
 DB 361 KALTIVGMLPAMFLVAGFGICDYSLAVAFITSTLTGSCSGSEINHLDIAPSYAGI 420
 QY 466 LIGITVTFATIPGVAVGVIAKSLTPDNTVGMQTVFYIAAIVFGAIFFTLPAKGEVON 525
 DB 421 LLSITVTFATIPGVAVGVIAKSLTPDNTVGMQTVFYIAAIVFGAIFFTLPAKGEVON 480
 QY 526 WALNDHHGH 534
 DB 481 MAIRTPSSH 489

RESULT 9
 ID 070580 ANOGA PRELIMINARY; PRT; 473 AA.
 AC 070580;
 DT 15-DEC-2003, integrated into UniProtKB/TrEMBL.
 DT 07-FEB-2006, entry version 2.
 DE ENSANGP0000010934 (Fragment).
 GN ORFNames=ENSANG00000008445;
 OS Anopheles gambiae str. PEST.
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culicidae;
 OC Anophelinae; Anopheles.
 NCBI_TaxID=180454;
 RX NUCLEOTIDE SEQUENCE.
 RP STRAIN=PEST.
 RG The Anopheles gambiae Sequence Committee;
 CC Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.

CC -----
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonCommercial License
 CC -----

DR EMBL; AAAB01008960; EAAL1677.2; -; Genomic DNA.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005215; F:transporter activity; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR007114; MFS.
 DR InterPro; IPR011701; MFS_1.
 DR Pfam; PF07690; MFS_1; 1.
 DR PROSITE; PS50850; MFS; 1.
 FT NON TER 1 1
 FT NON TER 473 473
 SQ SEQUENCE 473 AA; 52859 MM; FB6C0F77F35862A5 CRC64;

Query Match 37.1%; Score 1051; DB 2; Length 473;
 Best Local Similarity 44.6%; Pred. No. 3.4e-71;
 Matches 201; Conservative 89; Mismatches 143; Indels 18; Gaps 4;

QY 80 RYNLAIAFGFETVYALRVNLSVALVDMDSNTLTEDNRTSKACPEHSAPIKVHNQTK 139
 DB 10 RYIVFMAFLGFPNVVSLRVNLSVALVDMDSNTLTEDNRTSKACPEHSAPIKVHNQTK 55
 QY 140 KRYQMDAETQGMILGSPFYGYITQIPGVYASKIGKMLGFGILGTAVLTFTPIAAD 199
 DB 56 -YEDMSSLOGVYLSFFGYITLTPFGIGFSNRFGNRYFGVIGICTAVLTLLTPIAAK 114
 QY 200 LGVGPLVLRALBGLGEGVTFPAMHAMSSWAPPLERSKLLSISYAGOLGTVLSPLSG 259
 DB 115 AGVAVLAVRIVGIFEGVTFPCIHAWMSWAPPLERSKLLSISYAGOLGTVLSPLSG 174
 QY 260 IICVMMWTVVFFFGITGIFWFLMTWLVSDTPQKAKRISHKEKYLISLRNQSSOK 319
 DB 175 ILANAMGESVFFVFGIGIWMVAMFVKTSPEDVKWLSQREKELISLGRTEGEQ 234
 QY 320 SV--PWPPIKSLPLMAIVAHFSYNTFTLTLLPTYKEILRFVQENGFLSSLPYL 377
 DB 235 KIHGPKGIITSAVAVLAVSHSEKGFITLTQLPFLKDMHFELEKTVGSANPYL 294
 QY 438 TISTLTGFCSSGSPINHLDIAPSYAGILGINTPATIPGVAVGVIAKSLTPDNTVGM 497
 DB 354 TIAVGLGAFAMSGFAVHDLSPRSAGVLMGINSNTFATIPGIVSPITGYITNSKSD 413
 QY 498 QTVFYIAAIVFGAIFFTLPAKGEVONMAL 528
 DB 414 KTVFYIAAGIYLVGCVIYMFVGSDELQPM 444

RESULT 10
 ID 09YVG7 DROME PRELIMINARY; PRT; 559 AA.
 AC 09YVG7;
 DT 01-MAY-2000, integrated into UniProtKB/TrEMBL.
 DT 21-FEB-2006, entry version 1.
 DE CG4330-PA.
 GN ORFNames=CG4330; Dmel CG4330;
 OS Drosophila melanogaster (fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NCBI_TaxID=7227;
 RX NUCLEOTIDE SEQUENCE.
 RP MEDLINE=20196006; Pubmed=10731132; DOI=10.1126/science.287.5461.2185;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Suton G.G., Mortman J.R., Yandell M.D., Zhang O., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abiri J.F., Agbayan A., An H.-U., Andrews-Plamkoch C., Baldwin D.,
 RA Ballew R.M., Basu A.V., Bakendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borikova D., Botchan M.R., Bouck J., Brokstein P., Brooker P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahle C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrera S., Fleischmann W.,
 RA Fessler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glaeser K.,
 RA Glodok A., Gong F., Gottrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D.A., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasco P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskearn D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spralling A.C., Stapleton M., Strong R., Sun E.,
 RA Svitzkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasmann D.A., Weinstein G.M., Weisenbach J.,
 RA Williams S.M., Woodard T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang X., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RT "The genome sequence of *Drosophila melanogaster*." ,
 RT Science 287:2185-2195(2000) .
 [2]
 NUCLEOTIDE SEQUENCE.
 RA MEDLINE=22426065; PubMed=12537568;
 RP Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
 RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
 RA George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,
 RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
 RA Svitzkas R., Taber P.B., Wan K., Stapleton M., Sutton G.G., Venter C.,
 RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.,
 RT "Finishing a whole-genome shotgun: release 3 of the *Drosophila*
 RT *melanogaster* euchromatic genome sequence." ,
 RT Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002) .
 [3]
 NUCLEOTIDE SEQUENCE.
 RA MEDLINE=22426070; PubMed=12537573;
 RP Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J.W., Svitzkas R.,
 RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
 RA Ashburner M., Celniker S.E.,
 RT "The transposable elements of the *Drosophila melanogaster* euchromatin:
 RT a genomic perspective." ,
 RT Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002) .
 [4]
 NUCLEOTIDE SEQUENCE.
 RA MEDLINE=22426069; PubMed=12537572;
 RP Misra S., Crosby W.A., Mungall C.J., Matthews B.B., Campbell K.S.,
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
 RA Betencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 RA Lewis S.E.,
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a
 RT systematic review." ,
 RT Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002) .
 [5]
 NUCLEOTIDE SEQUENCE.
 RA Berkeley *Drosophila* Genome Project;
 RG Celniker S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,

RA Hoskins R., Stapleton M., Pacleb J., Park S., Svitzkas R., Smith E.,
 RA Yu C., Rubin G.,
 RT "Drosophila melanogaster release 4 sequence." ,
 RT Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
 [6]
 NUCLEOTIDE SEQUENCE.
 RA FlyBase;
 RP Submitted (JAN-2006) to the EMBL/Genbank/DBJ databases.
 CC -1- INTERACTION:
 CC 024180.Deaf1; NbrExp=1; IntAct=EBI-88239, EBI-149832;
 CC -1- SUBCELLULAR LOCATION: Membrane; multi-pass membrane protein (By
 CC similarity) .
 CC
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonDerivs license
 CC
 DR EMBL AB003491; AAF48230.1; -; Genomic_DNA.
 DR IntAct; Q9VYG7; -;
 DR FlyBase; FBgn0030452; CG4330.
 DR GO; GO:000515; F:protein binding; IPI.
 DR InterPro; IPR007114; MFS.
 DR InterPro; IPR011701; MFS_1.
 DR Pfam; PF07690; MFS_1; 1.
 DR PROSITE; PS50850; MFS; 1.
 KM Membrane; Transmembrane; Transport.
 SQ SEQUENCE 559 AA; 61745 MW; 8C48402881046C70 CRC64;
 Query Match 35.1%; Score 996.5; DB 2; Length 559;
 Best Local Similarity 39.8%; Pred. No. 5,7e-67;
 Matches 208; Conservative 91; Mismatches 171; Indels 53; Gaps 10;
 51 RNDGESTR-----TPLLPGAPPAEAAPVC-CSARYNALAIFGFFIYALRVNLS 102
 31 RSDDEADDERAFCSEGERLISSGAEEHNGCPTRHIFGMGLGFAVVYAMRVNLS 90
 103 VALVMDV-----SNTLEDNRTSKACEPHSAPIRVHNN-----QTGKTYQMDAETGG 150
 91 VALVAMVNOTALPHSNSVVIDTDT---CP-LPAP---HHNGSDPNPQKEGFVWDEATGG 143
 151 WLGSFFGYIITQIRGGVVASKIGKMLGFIIGTAVLTFTPIAADLGVPILVLA 210
 144 LVIGSFFGYVLTQVGGSMALVYGKKIYGVGLTAVFTLITPLAAMDLPVLVAVI 203
 211 LBSLGCVTFPAMHAAWSSAPPLERSKLLSISYAGAQGVYISLPLSGIIC---YWNW 267
 204 LBSMGCVTFPAMHAAWSSAPPLERSKLLSISYAGAQGVYISLPLSGIIC---YWNW 263
 268 TVVFFFGTIGFWEFLMIMLVSDTPQKRIISHYEKEYILSLRNQ----- 314
 264 PSAFYIFGLGILMFAMMYLVYDKPSDHPRISESEYIERSLOVORLINDLEABEE 323
 315 -----LSSQKSPWVPIKSLPLMAIVAHFYNMTFTLLTPTMYKEILRFNV 365
 324 EGODEVSLAPPEEPFIPMSLSLTVPLMILITOCQGMFAFTOLTLELFTYSNIHFPI 383
 366 QENGLFSLPYLGSWLCMLISGAPADNLRAKNFSTLCVRRIFSLIGMIGPAVLVAAGF 425
 384 QSNALLNAPVYLTWSVVGIALGALADMLARVYISLNSYKLMNTVASVPSGLIGITY 443
 426 ICGDYSLAVALTISTTLGGFCGSGSINHLDAEYAGILGITTTFATIPBMWCP-VI 484
 444 VCGDWMVTFMFLMGVGSFGGAVVAGQNMHIALSPYAGTMTGITSANITCGFLAPYVI 503
 485 AKSLTNDNTGEMQTFYIAAIVGALFPTLFAGEYQVNNAA 527
 504 GLIINRETLTQMHVFLAAGLINGNFTYILFASAEQSN 546
 RESULT 11
 OSVDMO_DROME PRELIMINARY; PRT; 502 AA.
 AC OSVDMO;
 DT 01-MAY-2000, integrated into UniProtKB/TrEMBL.

DT 01-MAY-2000, sequence version 1.
 DE 21-FEB-2006, entry version 33.
 DN CG4288-PA, isoform A (CG4288-pb, isoform b) (GH23975p).
 GN ORNAMES=CG4288, Dmel_CG4288;
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 CC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephyridiidae; Drosophilidae; Drosophila.
 NCBI_Taxid=7227;
 [1]
 RP NUCLEOTIDE SEQUENCE.
 RA MEDLINE=20196006; PubMed=10731132, DOI=10.1126/science.287.5461.2185;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.H.C., Blaise R.G., Champs E., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abail J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkov D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Dubin K.J., Evangelista C.C., Ferraz C., Ferlita S., Fleischmann W.,
 RA Fodor C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gottrell J.H., Gu Z., Guan F., Harris M.,
 RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
 RA Houlton D., Houston K.A., Howland T.J., Wei M.-H., Ibeagwam C.,
 RA Jaisl M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodali C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lako P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Mekulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Neidham D.R., Nelson K.A., Nixon R., Nuskern D.R., Paclab J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheller P., Shen H.,
 RA Shne B.C., Stiden-Klamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weissstock G.M., Weissbach J.,
 RA Williams S.M., Woodard T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 [2]
 RP NUCLEOTIDE SEQUENCE.
 RA MEDLINE=22426065; PubMed=12537568;
 RA Celniker S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,
 RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
 RA George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,
 RA Paclab J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
 RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
 RA Weissstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.,
 RT "Finishing a whole-genome shotgun: release 3 of the *Drosophila*
 RT *melanogaster* euchromatic genome sequence.";
 RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
 [3]
 RP NUCLEOTIDE SEQUENCE.
 RA MEDLINE=22426070; PubMed=12537573;
 RA Kamlirer J.S., Bergman C.M., Krommiller B., Carlson J.W., Svirskas R.,
 RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
 RA Ashburner M., Celniker S.E.;
 RT "The transposable elements of the *Drosophila melanogaster* euchromatin:
 RT a genomic perspective.";
 RL Genome Biol. 3:RESEARCH0084.1-RESEARCH0084.20(2002).
 [4]
 RP NUCLEOTIDE SEQUENCE.

RX MEDLINE=22426069; PubMed=12537572;
 RA Mista S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
 RA Hradecky P., Huang Y., Kaminker J.S., Milburn G.H., Prochuk S.E.,
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
 RA Betencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 RA Lewis S.E.;
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a
 RT systematic review.";
 RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
 [5]
 RP NUCLEOTIDE SEQUENCE.
 RA Berkeley *Drosophila* Genome Project;
 RA Celniker S., Carlson J., Wan K., Pfeiffer B., Frise E., George R.,
 RA Hoskins R., Stapleton M., Paclab J., Park S., Svirskas R., Smith E.,
 RA Yu C., Rubin G.;
 RT "Drosophila melanogaster release 4 sequence.";
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
 [6]
 RP NUCLEOTIDE SEQUENCE.
 RA Submitted (JAN-2006) to the EMBL/Genbank/DBJ databases.
 [7]
 RP NUCLEOTIDE SEQUENCE.
 RA STRAIN=Berkeley;
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
 RA Nuno J., Paclab J., Parasas V., Park S., Pounanavong S., Wan K.,
 RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
 RT Submitted (OCT-2001) to the EMBL/Genbank/DBJ databases.
 -I- SUPPLEMENTARY LOCATION: Membrane; multi-pass membrane protein (By
 similarity).

 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonCommercial License
 CC
 EMBL: AE003730; AAF55770.1; -; Genomic_DNA.
 EMBL: AY060776; AAL28324.1; -; mRNA.
 DR Inacac; Q9VDM0; -;
 DR Flybase; FBgn0038799; CG4288.
 DR GO; GO:0005515; F:protein binding; IPI.
 DR InterPro; IPR007114; MFS.
 DR InterPro; IPR011701; MFS_1.
 DR Pfam; PF07690; MFS_1; 1.
 DR PROSITE; PS50850; MFS; 1.
 KW Membrane; Transmembrane; Transport.
 KW SEQUENCE 502 AA; 5613 MW; 460D6FDDE741CE2 CRC64;
 SQ
 Query Match 35.0%; Score 992; DB 2; Length 502;
 Best Local Similarity 42.4%; Pred. No. 1.1e-66;
 Matches 190; Conservative 93; Mismatches 149; Indels 16; Gaps 3;
 QY 80 RYNAIAIAFFGFFIVYALRVNLVAVLDVMSNTTLLEDRTSKACPEHSAPIKYHHNDTG 139
 DB 38 RYIVVLAFFGFFVYVSVRLVLSVAIVAMENRIVFD-----AGGNSYQG-- 83
 QY 140 KKYQMDAETQWILGSPFFYGIITQIPGCVYASKIGCMGLGFGILGTAUVTLPPIAAD 199
 DB 84 -DFPMDSKOKALLISFFGYILTOFGIGYIGTIGXIVGCTGIGSAITLTLPMAAS 142
 QY 200 LGVGPFLVLRALDEGEGVTPPANAMWSSVAPLERSKLISVAGAQLTGIVSLPSG 259
 DB 143 HSLKMFLEKRIIEGFFBSVTPPGIHAVARNSPLESRWASINPAGYAGIVVAMPESG 202
 QY 260 IICVYMMVTYVFFGFTIGLFWFLMLWLTVSDTPQKHRIISHYKEYITLSLRNQLSQX 319
 DB 203 FLATKYGWESVFFYFGTIGVITWLVFVYAGBELDRFSGKECDYIKRTIGVGSCHV 262
 QY 320 SVFVNPILKSLPLMAIVVAHSYVWTRITLTLLPTWKELLRNVQENGSLSLPYIGS 379
 DB 263 KHPWRAIFTSMFPYAIMASHSESNWGFYITLTQPSFLRDTLNPDLGKTLISAVPYLAM 322

```

QY 380 WLCMLISGQAADNLRAKMNSTLCVRRISFLIGNIGPAVLVAAGFTICDYSLAVALTI 439
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 323 GILLAVSGYLADWLOVGIWTTQVRRNFNCGAFLAQTFPMMLTAYL-LDPTWSVSLTI 381
QY 440 STLLGFCSCSGFSINHLDIAPSYAGILGTINFPATIPGMVGPYIASLTPDNTVGEWOT 499
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 382 AVELGAFWMSGFVFNHLDIAPQHASVLMGIGNTFATIPGLIVSLTLGTIVTNTQSDENRI 441
QY 500 VFYIAAINVGAIFFTLFAKGEVQNMNA 527
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 442 IFRISAGIYVGCVIWYFCSGDLQEWNA 469
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 12
Q70579 ANOGA PRELIMINARY; PRT; 469 AA.
AC Q70579;
DT 15-DEC-2003, integrated into UniProtKB/TrEMBL.
DT 07-FEB-2004, sequence version 2.
DE ENSANGP0000011442 (Fragment).
GN ORFNames=ENSANGG0000008953;
OS Anopheles gambiae str. PEEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Preyigota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Culicidae;
OC Anophelinae; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=PEEST;
RG The Anopheles gambiae Sequence Committee;
RT "Anopheles gambiae re-annotation.";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
CC [-] CAUTION: The sequence shown here is derived from an
   EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
   preliminary data.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC
CC EMBL; AAB01008960; EAA11418.2; -; Genomic DNA.
CC GO: GO:0016021; C: integral to membrane; IEA.
CC GO: GO:0005215; P: transporter activity; IEA.
CC GO: GO:0006810; P: transport; IEA.
CC InterPro: IPR007114; MIPS.
CC InterPro: IPR011701; MIPS.
CC Pfam: PF07690; MIPS.1.
CC Prosite: PSS0850; MIPS.1.
FT NON_TER 1
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 469 AA; 52061 MW; BDBDA0AB09280936 CRC64;

Query Match 34.9%; Score 991; DB 2; Length 469;
Best local similarity 42.1%; Pred. No. 1.2e-66;
Matches 193; Conservative 93; Mismatches 154; Indels 18; Gaps 4;

QY 80 RYNLAIAFGGFIVYALRVNLVAALVMDVSNLTLENNRISKACPEHSAPIKYHANTG 139
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 8 RYLVFLAFAFGFVYVYSLRVNLVAIVAMTENRTVQYENGRT-----IGYEC-- 53
QY 140 KKYQWDAFTQCGWIIISGFFYYIIQIPGGYVASKIGGMLGFGILGTAVLTLPPIAAD 199
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 54 -EFWMDSTTKYIISFFYGIITQILGGYISNALGGYVGVGVGTAGLTILTPLAHH 112
QY 200 LGVGPLYLRALLEGISGVTFPAMHAMWSSWAPPLERSKLISYAGAQLGTIVSLPLSG 259
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 113 AGFWMLAVRATBEGFBEVTFPCIHAIWSNNAPEBSRMATIASGVFTGTIVASMLISG 172

```

```

QY 260 IICVYNNWTVYEFPGTIGIFMFLMIWLVSDTPQKHKISHYEKEYIISLRNQLSSQK 319
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 173 VLADLTGMEWVFITLILAFGLMFVAMMLIVKSPESDPITTKKFKFILLATLQSRADGGE 232
QY 320 SV--PWPVPIKSLPLMAIVVAHFSYNTFTLTLLPYWKEIILRFNVQENGLSLPYL 377
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 233 KYQHPRGRILTSKAVNALIVSSFSSEWNGFYTLTQPLTFLRDYMHELOMAGLSALPYL 292
QY 378 GSWLCLISGQAADNLRAKMNSTLCVRRISFLIGNIGPAVLVAAGFTICDYSLAVALTI 437
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 293 VMGLLSVAGYTLADLCQIRRMVLTQVRRFNCGAFLAQTFEVLGAYI-LRPAFLITCI 351
QY 438 TISTLLGFCSCSGFSINHLDIAPSYAGILGTINFPATIPGMVGPYIASLTPDNTVGEW 497
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 352 TIAVGCAPFAMCGFVFNHLDLSPKSAGVLMGISTFTSYAGILTPIVSQGLTASGSENEW 411
QY 498 QTVFYIAAINVGAIFFTLFAKGEVQNMALNDHHR 535
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 412 RTVFYIAAGIYLLIGCVTWFGVSGELQPMSEIERER 449

RESULT 13
YL2D CABEL STANDARD; PRT; 493 AA.
ID YL2D CABEL
AC 003567;
DT 01-FEB-1994, integrated into UniProtKB/Swiss-Prot.
DT 01-NOV-1997, sequence version 2.
DE 07-FEB-2006, entry version 49.
DE Hypothetical protein C38C10.2 in chromosome III.
GN ORFNames=C38C10.2;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Bristol N2;
RX MEDLINE=94150718; Pubmed=7906398; DOI=10.1038/368032a0;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berkis M.,
RA Bonfield J., Burton M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Frazer A.,
RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA Johnson L., Jones M., Kersey J., Kirtlen J., Laister N.,
RA Latreille P., Lighthouse J., Lloyd C., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkhen R.,
RA Sims M., Smaildon N., Smith A., Smith M., Sonhammer E., Stead R.,
RA Sulston J., Thierry-Mieg J., Thomas K., Vaubin M., Vaughan K.,
RA Waterston R., Watson A., Weinstock L., Wilkinson-Sproat J.,
RA Woldman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RT Nature 368:32-38 (1994).
RL Nature 368:32-38 (1994).
RN [2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; Pubmed=9851916; DOI=10.1126/science.282.5396.2012;
RG The C. elegans sequencing consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology.";
RL Science 282:2012-2018 (1998).
RN [3]
RP SEQUENCE REVISION.
RC Wormbase consortium;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
DB [-] SUBCELLULAR LOCATION: Membrane; multi-pass membrane protein
   (potential).
CC [-] SIMILARITY: Belongs to the major facilitator superfamily.
CC Sodium/anion cotransporter family.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC

```

DR EMBL; Z19153; CAA79549.1; -; Genomic DNA.
 DR PIR; G88553; G88553.
 DR Ensembl; C38C10.2; Caenorhabditis elegans.
 DR WormBase; WBGene00008000; C38C10.2.
 DR WormPep; C38C10.2; CE08647.
 DR InterPro; IPR007114; MFS.
 DR Pfam; PF07690; MFS_1; 1.
 DR PROSITE; PSS0850; MFS; 1.
 KM Complete proteome; Glycoprotein; Hypothetical protein; Ion transport;
 KM Membrane; Sodium; Sodium transport; Symport; Transmembrane; Transport.
 FT CHAIN 1 493
 FT
 FT TRANSMEM 10 30
 FT TRANSMEM 85 105
 FT TRANSMEM 112 132
 FT TRANSMEM 144 164
 FT TRANSMEM 175 195
 FT TRANSMEM 205 225
 FT TRANSMEM 272 292
 FT TRANSMEM 311 331
 FT TRANSMEM 348 368
 FT TRANSMEM 375 395
 FT TRANSMEM 406 426
 FT TRANSMEM 441 461
 FT BINDING 308 308
 FT
 FT CARBOHYD 35 35
 FT CARBOHYD 47 47
 FT CARBOHYD 69 69
 FT CARBOHYD 305 305
 FT CARBOHYD 433 433
 SQ SEQUENCE 493 AA; 52737 MW; 60792715D32553BD CRC64;
 Query Match 34.8%; Score 987; DB 1; Length 493;
 Best Local Similarity 39.7%; Pred. No. 2.6e-66;
 Matches 194; Conservative 94; Mismatches 179; Indels 22; Gaps 6;
 55 EESTDTPPLPGAPRAEAAPVCCSARNALIAFPFGFYIYARNLISVALVMDVSDNT 114
 2 EGATTPRLVP-----STRALSILVMEFGCLVYMKRTNNSFAVVCVMENKT 49
 115 LEDNRSKACPEHSAPIKVHNQTKKYQDAETOGWILCSFFGYITITPGGVASKI 174
 50 DTGVEKVRCKGKMTVESNSVIIG-EFMDKQTTGWNLSFFYGIIGSLASRY 108
 175 GGMALLGFGILGTAVALTLFTPIADLGVPLIVLRALGEGVTEPAMHMMSSVAPL 234
 109 GGRKVVFTILGSAALLTLNPAARTSEVALILRAIGFLQATFPAMHTMWSVNGPPL 168
 235 ESKSLISYAGQOLGTIVSLPSGLICYY--MMNTYVYFPGTIGIFFLIMLVSD 291
 169 ELSVLGVYVYAGQINIVILPLSGFCEYFGDGPSPSIFYIIGVFVLTAVWVYSSD 228
 292 TPQKHRIHYEKYLLSRNQLSSQ---KSPVWPILKSLPLMAIVVAHFSYMTFY 347
 229 KATHRIRITPEEKQYIVTAVASMGKDTGVSTPMIKILTSFAVWACMAHGACMDGAY 288
 348 TLTLTLPTYMKELIRENVQENGLSLPYLGSLWLCMLSGQADNLRAKNFSTLCVRI 407
 289 TMLVSLPSFLKDVLTGLNLSLGAVASIPYIAVFLAINAGVADTLRSKGLSTLTMRRA 348
 408 PSLIGMIGRAVFLVAAFGTIGCDYS-LAAVFLITSTLGGCCSGGSFSLNHLIDIPSTAGIL 466
 349 AMLVALNIGQIFLVASGYCGGDVVIIFITGMAISGQVAGFVNVYLEIAPPSGTV 408
 467 LGITNTEATIPGWWGVIAKSLTPDNTVGEVQTVFIATAINVGAIFFTLPAKGEQNW 526
 409 MGTGNTISALAGIISPAVSSYLTNPQTQEBMWVLTAGILITGALLPSIFASGEVQPM 468
 527 A-LNDHGH 534

DB 469 AKLTAERGH 477
 RESULT 14
 ID Q920B7_MOUSE PRELIMINARY; PRT; 582 AA.
 AC Q920B7;
 DT 01-DEC-2001, integrated into UniProtKB/TrEMBL.
 DT 01-DEC-2001, sequence version 1.
 DT 07-FEB-2006, entry version 25.
 DE Vesicular glutamate transporter 2 (Solute carrier family 17 (sodium-dependent inorganic phosphate cotransporter), member 6).
 GN Names=SLC17a6;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RX MEDLINE=21453279; PubMed=11432869; DOI=10.1074/jbc.M104578200;
 RA Bai L., Xu H., Collins J.F., Ghishan F.K.;
 RT "Molecular and functional analysis of a novel neuronal vesicular
 glutamate transporter.";
 RL J. Biol. Chem. 276:36764-36769 (2001).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Eye;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ueidl T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Muliyil S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywicki M.I., Skalek U., Smilun D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Merra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Eye;
 RG NIH MGC Project;
 RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>
 CC Distributed under the Creative Commons Attribution-NonDerivs license

EMBL; AF324864; AAL08941.1; -; mRNA.
 EMBL; BC038375; AAH38375.1; -; mRNA.
 DR Ensembl; ENSMUSG00000030500; Mus musculus.
 MGI; MGI:2156052; Slc17a6.
 DR GO; GO:0016021; C:integral to membrane; RCA.
 DR GO; GO:0016020; C:membrane; RCA.
 DR GO; GO:0008021; C:synaptic vesicle; IDA.
 DR GO; GO:0005113; F:1-glutamate transporter activity; IDA.
 DR GO; GO:0005215; F:transporter activity; RCA.
 DR GO; GO:0001504; P:neurotransmitter uptake; IDA.
 DR GO; GO:0006810; P:transport; RCA.
 DR InterPro; IPR007114; MFS.
 DR Pfam; PF07690; MFS_1; 1.
 DR PROSITE; PSS0850; MFS; 1.

KW Membrane: Transmembrane; Transport.
 SQ SEQUENCE 582 AA; 64560 MW; 9F7A4F62E6859A8AD CRC64;

Query Match 34.2%; Score 970; DB 2; Length 582;

Best Local Similarity 40.9%; Pred. No. 6, 2e-65; Matches 201; Conservative 93; Mismatches 165; Indels 32; Gaps 9;

```

QY 51 RNDGEESTRTPLPCAP---RAEAPVC-CSA-----RYNLAIAFGFFIYALRVNL 101
DB 34 KQDNRETIILTE--DCKPLEVEPKKAPLDCCTCFGLPRRYIIMSGLFCISFGIRCNL 91
QY 102 SYALVDMVDSNTTLEDNRTSKACPEHSAPIKVHHNOTGKKYQWDAETQGMILGSPFYGY 161
DB 92 GVAIVDMVNSSTI---HRGKVIKEKA-----KENWDPETVGMHGSFFWGYI 136
QY 162 ITQIPGGYVASKIGKMLGFGILGTAUVTLPPIAADLGVGPLYLALREGLEGVTPP 221
DB 137 ITQIPGGYIASRLAANRVFGAAILLTSTLMMLPSAARVHYGCVIFVRILQGLVBSGYTP 196
QY 222 ANHAMSSWAPPLERSKLTISYAGAQLTGVTISLPSGIIICYMMNTYVFFFGTIGIFW 281
DB 197 ACHGISKWAAPPLERSRLATTSFCGSYAGAVIAMPLAGILVQGTGSSVFFYVYGSFGMW 256
QY 282 FLIMTWLVSPTQKHKRISHYEKEYILSL--RNQLSQK--SVWPVILKSLPLMAIV 336
DB 257 YMFWLVSYESPAKHPTITDEERRYIEESIGESANILGAMKEKFTPMRKFTSMPLYAII 316
QY 337 VAHFSYNMTFFYLLTLPTMYKEILRFNVOENGFLSLPYLGSWLCMLISGQAADNLRAK 396
DB 317 VANFCSSWTFYLLISQPAFFEEVGFELSKVMTLAVPLVMTIIVPIGGQIADPLRSK 376
QY 397 WNFSTLCVRRIFFSLIGMIGPAVFLVAAFGICDYSILAVALFTISTTGGFCSSGFSINHL 456
DB 377 QILSTTVYKINMCGFGMEATLLLVGV-SHTRGVAISFLVLAVGSGFAISGRNVNL 435
QY 457 DIAPSYAGILGLITNTFATIPGMVGPVIAKSLTPDNTVGEMQTVFYIAAIVFGAIFPT 516
DB 436 DIAPRYASITLMGISNGVLTSGMVCPIVGAAMTKNSREBQYVFLIALVHYGVIFYA 495
QY 517 LFAKGEVQWMA 527
DB 496 LFASGEKQPMW 506

```

RESULT 15

ID Q9J112_RAT PRELIMINARY; PRT; 582 AA.

AC Q9J112; 01-OCT-2000, integrated into UniProtKB/TrEMBL.

DT 01-OCT-2000, sequence version 1.

DT 07-FEB-2006, entry version 20.

DE Differentiation-associated Na-dependent inorganic phosphate

DE cotransporter.

DE Name=Slc17a6; Synonym=DNP1;

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;

OC Muridae; Muridae; Murinae; Rattus.

OC NCBI_TaxId=10116;

RN (1)

RP NUCLEOTIDE SEQUENCE.

RA Mashima H., Kojima I.;

RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.

CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>

CC Distributed under the Creative Commons Attribution-NonDerivs License

CC EMBL; AF271235; AAF76223.1; -; mRNA.

DR Ensembl: ENSRNOG0000016147; Rattus norvegicus.

DR RGD: 620531; Slc17a6.

DR GO: GO:0005769; C:early endosome; IDA.

DR GO: GO:0008021; C:synaptic vesicle; TAS.

DR GO: GO:0005313; F:L-glutamate transporter activity; IDA.

DR GO: GO:0001504; P:neurotransmitter uptake; TAS.

DR InterPro: IPR007114; MFS.

DR InterPro: IPR011701; MFS_1.

DR Pfam: PF07690; MFS_1.

DR PROSITE: PSS0850; MFS_1.

KW Membrane; Transmembrane; Transport.

SQ SEQUENCE 582 AA; 64576 MW; 99A14F62E6859E9 CRC64;

Query Match 34.2%; Score 970; DB 2; Length 582;

Best Local Similarity 40.9%; Pred. No. 6, 2e-65; Matches 201; Conservative 93; Mismatches 165; Indels 32; Gaps 9;

```

QY 51 RNDGEESTRTPLPCAP---RAEAPVC-CSA-----RYNLAIAFGFFIYALRVNL 101
DB 34 KQDNRETIILTE--DCKPLEVEPKKAPLDCCTCFGLPRRYIIMSGLFCISFGIRCNL 91
QY 102 SYALVDMVDSNTTLEDNRTSKACPEHSAPIKVHHNOTGKKYQWDAETQGMILGSPFYGY 161
DB 92 GVAIVDMVNSSTI---HRGKVIKEKA-----KENWDPETVGMHGSFFWGYI 136
QY 162 ITQIPGGYVASKIGKMLGFGILGTAUVTLPPIAADLGVGPLYLALREGLEGVTPP 221
DB 137 ITQIPGGYIASRLAANRVFGAAILLTSTLMMLPSAARVHYGCVIFVRILQGLVBSGYTP 196
QY 222 ANHAMSSWAPPLERSKLTISYAGAQLTGVTISLPSGIIICYMMNTYVFFFGTIGIFW 281
DB 197 ACHGISKWAAPPLERSRLATTSFCGSYAGAVIAMPLAGILVQGTGSSVFFYVYGSFGMW 256
QY 282 FLIMTWLVSPTQKHKRISHYEKEYILSL--RNQLSQK--SVWPVILKSLPLMAIV 336
DB 257 YMFWLVSYESPAKHPTITDEERRYIEESIGESANILGAMKEKFTPMRKFTSMPLYAII 316
QY 337 VAHFSYNMTFFYLLTLPTMYKEILRFNVOENGFLSLPYLGSWLCMLISGQAADNLRAK 396
DB 317 VANFCSSWTFYLLISQPAFFEEVGFELSKVMTLAVPLVMTIIVPIGGQIADPLRSK 376
QY 397 WNFSTLCVRRIFFSLIGMIGPAVFLVAAFGICDYSILAVALFTISTTGGFCSSGFSINHL 456
DB 377 QILSTTVYKINMCGFGMEATLLLVGV-SHTRGVAISFLVLAVGSGFAISGRNVNL 435
QY 457 DIAPSYAGILGLITNTFATIPGMVGPVIAKSLTPDNTVGEMQTVFYIAAIVFGAIFPT 516
DB 436 DIAPRYASITLMGISNGVLTSGMVCPIVGAAMTKNSREBQYVFLIALVHYGVIFYA 495
QY 517 LFAKGEVQWMA 527
DB 496 LFASGEKQPMW 506

```

Search completed: June 7, 2006, 05:42:53

Job time : 304 secs

FT	/label= PKC_phospho_site
FT	/note= "Putative phosphorylation site"
FT	Modified-site
FT	136. .139
FT	/note= "Asn is N-glycosylated"
FT	Modified-site
FT	138. .141
FT	/note= "Putative amidation site"
FT	Modified-site
FT	138. .140
FT	/label= PKC_phospho_site
FT	/note= "Putative phosphorylation site"
FT	Modified-site
FT	167. .172
FT	/note= "Putative myristylation site"
FT	Modified-site
FT	183. .188
FT	/note= "Putative myristylation site"
FT	Modified-site
FT	213. .218
FT	/note= "Putative myristylation site"
FT	Modified-site
FT	246. .251
FT	/note= "Putative myristylation site"
FT	Modified-site
FT	250. .255
FT	/note= "Putative myristylation site"
FT	Modified-site
FT	266. .269
FT	/note= "Asn is N-glycosylated"
FT	Modified-site
FT	297. .300
FT	/label= CAMP_phospho_site
FT	/note= "Putative phosphorylation site"
FT	Modified-site
FT	300. .303
FT	/label= CK2_phospho_site
FT	/note= "Putative phosphorylation site"
FT	Modified-site
FT	310. .312
FT	/label= PKC_phospho_site
FT	/note= "Putative phosphorylation site"
FT	Modified-site
FT	317. .320
FT	/label= PKC_phospho_site
FT	/note= "Putative phosphorylation site"
FT	Modified-site
FT	333. .346
FT	/note= "Asn is N-glycosylated"
FT	Modified-site
FT	378. .383
FT	/note= "Putative myristylation site"
FT	Modified-site
FT	398. .401
FT	/note= "Asn is N-glycosylated"
FT	Modified-site
FT	427. .432
FT	/note= "Putative myristylation site"
FT	Modified-site
FT	444. .449
FT	/note= "Putative myristylation site"
FT	Modified-site
FT	464. .469
FT	/note= "Putative myristylation site"
FT	Modified-site
FT	468. .473
FT	/note= "Putative myristylation site"
FT	Modified-site
FT	493. .496
FT	/label= CK2_phospho_site
FT	/note= "Putative phosphorylation site"
PN	WO200005375-A1.
XX	03-FEB-2000.
XX	
XX	22-JUL-1999;
XX	99WO-US016676.
XX	22-JUL-1998;
XX	98US-0093843P.
XX	(UYVA-) UNIV VANDERBILT.
XX	
XX	HELLERVIET CG, FU C;
XX	
XX	WPI; 2000-205377/18.
XX	DR N-PSDB; AAZ50879.
XX	
XX	New polynucleotide encoding mammalian receptor for streptococcus toxin,
XX	useful for diagnosis and treatment of, e.g. pneumonia in neonates.
XX	Claim 10; Page 93-95; 109pp; English.
XX	
XX	The present sequence is partial human GBS (group B beta-haemolytic
XX	streptococci) toxin receptor (HP55). GBS toxin receptor is an integral

[illegible]

```

OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH Region 8..28
FT Region /label= Hab3_immunogenic_peptide
FT Region /label= Hab4_immunogenic_peptide
FT Region /label= Hab4_immunogenic_peptide
FT Region 49..76
FT Region /label= Hab4_immunogenic_peptide
FT Region 49..63
FT Region /label= Hab1_immunogenic_peptide
FT Region 112..125
FT Region /label= Hab2_immunogenic_peptide
XX
XX MO200156598-A2.
XX
XX PD 09-AUG-2001.
XX
XX PF 02-FEB-2001; 2001WO-US003662.
XX
XX PR 02-FEB-2000; 2000US-0179870P.
XX
XX PA (UYVA-) UNIV VANDERBILT.
XX
XX PI Hellerqvist CG;
XX
XX DR WPI; 2001-488844/53.
XX
XX DR N-PSDB; AAD10325.
XX
XX PT Preventing or attenuating pathoangiogenic conditions e.g. cancer, chronic
XX PT wounds, osteoarthritis, keloids and psoriasis in a mammal, by
XX PT administering group B beta-hemolytic streptococci toxin receptor or its
XX PT fragment.
XX
XX PS Claim 4; Page 45-47; 52pp; English.
XX
XX CC The present sequence is full length group B beta-haemolytic Streptococci
XX CC (GBS) toxin receptor protein, HP59 from human. The present invention
XX CC relates to a method for preventing or attenuating a patho-angiogenic
XX CC condition in a mammal which comprises administering to the mammal one or
XX CC more GBS toxin receptors or their immunogenic fragments to induce or
XX CC maintain an immune response to one of GBS toxin receptors. The method is
XX CC useful for preventing or ameliorating pathoangiogenic conditions such as
XX CC cancer, scarring during wound healing, gliosis during repair of nerve
XX CC injury, chronic wounds, keloids, reperfusion injury, rheumatoid
XX CC arthritis, atherosclerosis, osteoarthritis and psoriasis in mammals. The
XX CC proteins of the invention are also used as vaccines
XX
XX SQ Sequence 536 AA;
XX
XX Query Match 100.0%; Score 2836; DB 4; Length 536;
XX Best Local Similarity 100.0%; Pred. No. 1.8e-282;
XX Matches 536; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 MAAGAMTPRPVOPARPGFGLSGRSLCCVASTPAHVGMRSVPVDLARNDEESTDR 60
XX |
XX |
XX |
XX 1 MAAGAMTPRPVOPARPGFGLSGRSLCCVASTPAHVGMRSVPVDLARNDEESTDR 60
XX |
XX |
XX |
XX DB 61 TPLLPGARPAEAAPVCCSARYNLAILAFPGFITYAARVNLVAVLVMVDSNTTLEENRT 120
XX |
XX |
XX |
XX DB 61 TPLLPGARPAEAAPVCCSARYNLAILAFPGFITYAARVNLVAVLVMVDSNTTLEENRT 120
XX |
XX |
XX |
XX QY 121 SKACPEHSAPIKVHNQTKKKYQMDATOGWILGSPFYGYITITIPGQYVASKIGKML 180
XX |
XX |
XX |
XX DB 121 SKACPEHSAPIKVHNQTKKKYQMDATOGWILGSPFYGYITITIPGQYVASKIGKML 180
XX |
XX |
XX |
XX QY 121 SKACPEHSAPIKVHNQTKKKYQMDATOGWILGSPFYGYITITIPGQYVASKIGKML 180
XX |
XX |
XX |
XX DB 121 SKACPEHSAPIKVHNQTKKKYQMDATOGWILGSPFYGYITITIPGQYVASKIGKML 180
XX |
XX |
XX |
XX QY 181 GFGILGTAIVLTLPPIADLGVPLIVLRALBGLGEGVTPAMHAMSSNAPPLERSKLT 240
XX |
XX |
XX |
XX DB 181 GFGILGTAIVLTLPPIADLGVPLIVLRALBGLGEGVTPAMHAMSSNAPPLERSKLT 240
XX |
XX |
XX |
XX QY 241 SISVAGNOLGTVISLPSGIIICYMMNTYVYFPGTIGIFPFLIMLVSTPQKHRIIS 300
XX |
XX |
XX |
XX DB 241 SISVAGNOLGTVISLPSGIIICYMMNTYVYFPGTIGIFPFLIMLVSTPQKHRIIS 300
XX |
XX |
XX |
XX QY 301 HYEKEYILSLRNQLSSQKSVFWVPIILKSILPMLAIVVAHFSYNNTFYTLTLPTVMKEI 360
XX |
XX |
XX |

```

```

DB 301 HYEKEYILSLRNQLSSQKSVFWVPIILKSILPMLAIVVAHFSYNNTFYTLTLPTVMKEI 360
XX |
XX |
XX |
QY 361 LRFVVOENGFLSSIPYIGSWLCTMLSGAADNLRANKNFSTLCYRRIFSLIGMTGPAVFL 420
XX |
XX |
XX |
DB 361 LRFVVOENGFLSSIPYIGSWLCTMLSGAADNLRANKNFSTLCYRRIFSLIGMTGPAVFL 420
XX |
XX |
XX |
QY 421 VAAFGICDYSILAVAFITISTTGGFCCSGFSINHLDIAPSYAGILGINTPATIPGMV 480
XX |
XX |
XX |
DB 421 VAAFGICDYSILAVAFITISTTGGFCCSGFSINHLDIAPSYAGILGINTPATIPGMV 480
XX |
XX |
XX |
QY 481 GPVIAKSLTPDNTVGEWQTVFYIAAIVFGAIFFTLPKAGEVONMALNDHGHRRH 536
XX |
XX |
XX |
DB 481 GPVIAKSLTPDNTVGEWQTVFYIAAIVFGAIFFTLPKAGEVONMALNDHGHRRH 536
XX |
XX |
XX |
RESULT 3
AAY45087
ID AAY45087 standard; protein; 495 AA.
XX
XX AC AAY45087;
XX
XX DT 31-MAY-2000 (first entry)
XX
XX DE Partial human GBS toxin receptor (HP55) .
XX
XX KW Human GBS toxin receptor; group B beta-haemolytic streptococci; HP55;
XX KW pathological vascularisation; cancer metastases; angiogenesis;
XX KW neovascularisation; reperfusion injury; scarring; keloid;
XX KW chronic inflammation; rheumatoid arthritis; psoriasis; neural injury;
XX KW endothelial cell proliferation; antibacterial; anticancer;
XX KW anti-angiogenic; anti-inflammatory; anti-arthritis; anti-psoriatic.
XX
XX OS Homo sapiens.
XX
XX CC Key Location/Qualifiers
XX CC FH 14..20
XX CC FT /note= "Region of high hydrophilicity used to design
XX CC FT antigenic peptide"
XX CC FT 26..30
XX CC FT /note= "Region of high hydrophilicity used to design
XX CC FT antigenic peptide"
XX CC FT 76..80
XX CC FT /note= "Region of high hydrophilicity used to design
XX CC FT antigenic peptide"
XX
XX PS WO200005375-A1.
XX
XX PN 03-FEB-2000.
XX
XX PD 22-JUL-1999; 99WO-US016676.
XX
XX PR 22-JUL-1998; 98US-0093843P.
XX
XX PA (UYVA-) UNIV VANDERBILT.
XX
XX PI Hellerqvist CG, Fu C;
XX
XX DR WPI; 2000-205377/18.
XX
XX DR N-PSDB; AAZ50875.
XX
XX PT New polynucleotide encoding mammalian receptor for streptococcus toxin,
XX PT useful for diagnosis and treatment of, e.g. pneumonia in neonates.
XX
XX PS Claim 9; Page 22; 109pp; English.
XX
XX CC The present sequence is partial human GBS (group B beta-haemolytic
XX CC streptococci) toxin receptor (HP55). GBS toxin receptor is an integral
XX CC protein with seven transmembrane domains. Expression vectors comprising
XX CC the coding region can be transformed into host cells to express GBS toxin
XX CC receptor and its fragments. Detecting the receptor in tissues is used to
XX CC diagnose pathological vascularisation, e.g. for detecting cancer
XX CC metastases. GBS toxin receptors are useful for treating conditions

```

CC associated with pathological angiogenesis or neovascularisation
CC (specifically cancer, reperfusion injury, scarring during wound healing,
CC keloids, chronic inflammation (rheumatoid arthritis or psoriasis) or
CC neural injury), and to raise specific antibodies used for treating early
CC onset disease. Inhibitors of this receptor are useful for treating
CC pathological or hypoxia-induced endothelial cell proliferation and
CC migration

XX
SQ Sequence 495 AA;

Query Match 92.4%; Score 2621; DB 3; Length 495;

Best Local Similarity 100.0%; Pred. No. 2,2e-260;

Matches 495; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 MRSPPVRLARNDESESTDRTPLLPGAPRAEAPVCCSARYNLALAFEGFFIVALARVNL 101
DB 1 MRSPPVRLARNDESESTDRTPLLPGAPRAEAPVCCSARYNLALAFEGFFIVALARVNL 60
QY 102 SVALVDMVDSNTLTEDNRTSKACEHSAPIKVHNQTKKYQMDAETQGMILGSFFGYI 161
DB 61 SVALVDMVDSNTLTEDNRTSKACEHSAPIKVHNQTKKYQMDAETQGMILGSFFGYI 120
QY 162 ITQIPGGYVASKIGKMLGFGILGTAVALTLFTPIAADLGVGPLIVRALBEGLEGVTFP 221
DB 121 ITQIPGGYVASKIGKMLGFGILGTAVALTLFTPIAADLGVGPLIVRALBEGLEGVTFP 180
QY 222 AMHAMSSWAPPLERSKLSISYAGQGTIVISLPLSGIICYYMMNTYVFFFGTIGIFW 281
DB 181 AMHAMSSWAPPLERSKLSISYAGQGTIVISLPLSGIICYYMMNTYVFFFGTIGIFW 240
QY 282 FLIMWLVSPTPOKHRIKISHYEKEYILSLRNQSSOKSVPMVPIKSLPLMAIVVAHFS 341
DB 241 FLIMWLVSPTPOKHRIKISHYEKEYILSLRNQSSOKSVPMVPIKSLPLMAIVVAHFS 300
QY 342 YNMTFYTLTLPTVYKKEILRFVQENGFLSSLPYLGSWLCMILSGQAADNLRKWNFST 401
DB 301 YNMTFYTLTLPTVYKKEILRFVQENGFLSSLPYLGSWLCMILSGQAADNLRKWNFST 360
QY 402 LCVRRIFFSLIGMIGPAVFLVAAFGICDYSLVAFLITITLGGFCSSGFSINHLDIAPS 461
DB 361 LCVRRIFFSLIGMIGPAVFLVAAFGICDYSLVAFLITITLGGFCSSGFSINHLDIAPS 420
QY 462 YAGILGINTNTFATIPGMVGPVIAKSLTPDNTVGEMQTVFYIAAIVNFGAIFFTLPFAKG 521
DB 421 YAGILGINTNTFATIPGMVGPVIAKSLTPDNTVGEMQTVFYIAAIVNFGAIFFTLPFAKG 480
QY 522 EVQNMALNDHHGHRH 536
DB 481 EVQNMALNDHHGHRH 495

RESULT 4

ID AAB66967 standard; protein; 495 AA.

AC AAB66967;

DT 18-APR-2001 (first entry)

DE Human AST.

XX Human; AST; noctropic; immunotropic; gene therapy; Salla disease;

KM anion and sugar transporter; anion-cation symporter;

XX Human; AST; noctropic; immunotropic; gene therapy; Salla disease;

XX Human; AST; noctropic; immunotropic; gene therapy; Salla disease;

XX Human; AST; noctropic; immunotropic; gene therapy; Salla disease;

XX Human; AST; noctropic; immunotropic; gene therapy; Salla disease;

PF 16-JUL-1999; 99EP-00202341.

PR 16-JUL-1999; 99EP-00202341.

XX (ALKU) AKZO NOBEL NV.

XX WPI; 2001-193090/20.

DR N-PSDB; AAF55900.

XX New human transporter gene implicated in Salla disease and lysosomal

PT sialic acid transport, useful in assays for identifying new drugs, or

PT diagnosing sialic acid transport defects related to mutations in the

PT transporter gene.

XX Claim 1; Page 10-12; 20pp; English.

XX The present sequence is human Anion and Sugar Transporter (AST) protein

CC sequence. AST has significant homology with several members of the anion-

CC cation symporter (ACS) family of transporters. AST is implicated in Salla

CC disease, and is useful in screening assays for identifying new drugs.

CC Compounds identified via AST screening is useful for preparing a

CC pharmaceutical suitable as an activator or inhibitor of a sialic acid

CC transporter protein. The pharmaceutical may be used in sialic acid

CC associated diseases and CNS/immune related disorders

XX Sequence 495 AA;

Query Match 92.4%; Score 2621; DB 4; Length 495;

Best Local Similarity 100.0%; Pred. No. 2,2e-260;

Matches 495; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 MRSPPVRLARNDESESTDRTPLLPGAPRAEAPVCCSARYNLALAFEGFFIVALARVNL 101
DB 1 MRSPPVRLARNDESESTDRTPLLPGAPRAEAPVCCSARYNLALAFEGFFIVALARVNL 60
QY 102 SVALVDMVDSNTLTEDNRTSKACEHSAPIKVHNQTKKYQMDAETQGMILGSFFGYI 161
DB 61 SVALVDMVDSNTLTEDNRTSKACEHSAPIKVHNQTKKYQMDAETQGMILGSFFGYI 120
QY 162 ITQIPGGYVASKIGKMLGFGILGTAVALTLFTPIAADLGVGPLIVRALBEGLEGVTFP 221
DB 121 ITQIPGGYVASKIGKMLGFGILGTAVALTLFTPIAADLGVGPLIVRALBEGLEGVTFP 180
QY 222 AMHAMSSWAPPLERSKLSISYAGQGTIVISLPLSGIICYYMMNTYVFFFGTIGIFW 281
DB 181 AMHAMSSWAPPLERSKLSISYAGQGTIVISLPLSGIICYYMMNTYVFFFGTIGIFW 240
QY 282 FLIMWLVSPTPOKHRIKISHYEKEYILSLRNQSSOKSVPMVPIKSLPLMAIVVAHFS 341
DB 241 FLIMWLVSPTPOKHRIKISHYEKEYILSLRNQSSOKSVPMVPIKSLPLMAIVVAHFS 300
QY 342 YNMTFYTLTLPTVYKKEILRFVQENGFLSSLPYLGSWLCMILSGQAADNLRKWNFST 401
DB 301 YNMTFYTLTLPTVYKKEILRFVQENGFLSSLPYLGSWLCMILSGQAADNLRKWNFST 360
QY 402 LCVRRIFFSLIGMIGPAVFLVAAFGICDYSLVAFLITITLGGFCSSGFSINHLDIAPS 461
DB 361 LCVRRIFFSLIGMIGPAVFLVAAFGICDYSLVAFLITITLGGFCSSGFSINHLDIAPS 420
QY 462 YAGILGINTNTFATIPGMVGPVIAKSLTPDNTVGEMQTVFYIAAIVNFGAIFFTLPFAKG 521
DB 421 YAGILGINTNTFATIPGMVGPVIAKSLTPDNTVGEMQTVFYIAAIVNFGAIFFTLPFAKG 480
QY 522 EVQNMALNDHHGHRH 536
DB 481 EVQNMALNDHHGHRH 495

RESULT 5

ID ADJ75516 standard; protein; 495 AA.

AC ADJ75516;

DT 20-MAY-2004 (first entry)

XX Marker gene related amino acid sequence SEQ ID NO:768.
 DE bronchial asthma; chronic obstructive pulmonary disease;
 KM respiratory epithelial cell; interleukin-13; respiratory; antiasthmatic;
 KM gene therapy; marker.
 XX Homo sapiens.
 OS
 XX
 PN EP1394274-A2.
 XX
 PD 03-MAR-2004.
 XX
 PF 04-AUG-2003; 2003EP-00254857.
 XX
 PR 06-AUG-2002; 2002JP-00229312.
 PR 20-MAR-2003; 2003JP-00077212.
 XX
 PA (GENO-) GENOX RES INC.
 XX
 PI Ohtani N, Sugita Y, Yamaya M, Kubo H, Nagai H, Izuhara K;
 XX WPI; 2004-193155/19.
 DR
 XX
 PT Testing for bronchial asthma or chronic obstructive pulmonary disease by
 PT comparing the expression level of a marker gene in a biological sample
 PT from a subject with the expression level of the gene in a sample from a
 PT healthy subject.
 XX
 PS Example 11; SEQ ID NO 768; 241pp; English.
 XX
 CC The present invention describes a method of testing for bronchial asthma
 CC or chronic obstructive pulmonary disease. The method comprises
 CC determining the expression level of a marker gene in a biological sample
 CC from a subject, comparing the expression level determined with the
 CC expression level of the marker gene in a biological sample from a healthy
 CC subject, and judging whether the subject has bronchial asthma or chronic
 CC obstructive pulmonary disease. The marker gene comprises: (a) a group of
 CC genes (S1) whose expression levels increase when respiratory epithelial
 CC cells are stimulated with interleukin-13; or (b) a group of genes (S2)
 CC whose expression levels decrease when respiratory epithelial cells are
 CC stimulated with interleukin-13. Also described: (1) a reagent (I) for
 CC testing for bronchial asthma or chronic obstructive pulmonary disease;
 CC (2) a kit for screening for a candidate compound for a therapeutic agent
 CC to treat bronchial asthma or chronic obstructive pulmonary disease; (3)
 CC an animal model for bronchial asthma or chronic obstructive pulmonary
 CC disease; (4) an inducer that induces bronchial asthma in a mouse; (5) a
 CC method for producing an animal model for bronchial asthma or chronic
 CC obstructive pulmonary disease; (6) a therapeutic agent for bronchial
 CC asthma or chronic obstructive pulmonary disease, comprising the compound,
 CC a marker gene or an antisense nucleic acid corresponding to a portion of
 CC the marker gene, a ribozyme, a polynucleotide that suppresses the
 CC expression of the gene through an RNAi effect or an antibody recognising
 CC a protein encoded by a marker gene; and (7) a DNA chip for testing for
 CC bronchial asthma or a chronic obstructive pulmonary disease, on which a
 CC probe has been immobilised to assay a marker gene. (1) has respiratory
 CC and antiasthmatic activities, and can be used in gene therapy. The method
 CC is useful for testing for or screening for a therapeutic agent for
 CC bronchial asthma or chronic obstructive pulmonary disease. The present
 CC sequence is used in the exemplification of the present invention.
 XX
 SQ Sequence 495 AA:
 Query Match 92.4%; Score 2621; DB 8; Length 495;
 Best Local Similarity 100.0%; Pred. No. 2,2e-260; Mismatches 0; Gaps 0;
 Matches 495; Conservative 0; Indels 0;
 QY 42 MRSPVRLARNDGSESTDRPTLLPGAPRAEAPVCCSARYVLTALAFGFIYALRNLT 101
 DB 1 MRSVRLARNDGSESTDRPTLLPGAPRAEAPVCCSARYVLTALAFGFIYALRNLT 60
 QY 102 SVALVDMVDSNTLTEDNRTSKACEHSAPIKVHNQTKKYYQMDAETQGMILGSFFGYI 161

DB 61 SVALVDMVDSNTLTEDNRTSKACEHSAPIKVHNQTKKYYQMDAETQGMILGSFFGYI 120
 QY 162 ITQIPGVASKIGKGMILGFILGTAVLTLPFIADLGVGPILVLRALGEGVTFP 221
 DB 121 ITQIPGVASKIGKGMILGFILGTAVLTLPFIADLGVGPILVLRALGEGVTFP 180
 QY 222 AMHAMSSWAPPLERSKILSISVGAOLGVISLPLSGIICYNMWTVVFFFTIGIFW 281
 DB 181 AMHAMSSWAPPLERSKILSISVGAOLGVISLPLSGIICYNMWTVVFFFTIGIFW 240
 QY 282 FLIMTWLVSPTPOGKRIISHEKEYIISLNLQSSQKSVWVPLILSLPLMAIVVAHS 341
 DB 241 FLIMTWLVSPTPOGKRIISHEKEYIISLNLQSSQKSVWVPLILSLPLMAIVVAHS 300
 QY 342 YNMFTFYTLTLPLPYMKELIFENVOENGFLSSLPYLGSWLCMISGQADNLRKMNST 401
 DB 301 YNMFTFYTLTLPLPYMKELIFENVOENGFLSSLPYLGSWLCMISGQADNLRKMNST 360
 QY 402 LCVRRIFSLIGMIGPAVFLVAGFICGDDYSIAVAFITISTTLGFCSSGFSINHLIAPS 461
 DB 361 LCVRRIFSLIGMIGPAVFLVAGFICGDDYSIAVAFITISTTLGFCSSGFSINHLIAPS 420
 QY 462 YAGILLGITNTFATITPGWGVPIAKSLTPDNTVGSEMOVFYIAAINVFGAIFPTLPFKG 521
 DB 421 YAGILLGITNTFATITPGWGVPIAKSLTPDNTVGSEMOVFYIAAINVFGAIFPTLPFKG 480
 QY 522 EVQNWALNDHGHHRH 536
 DB 481 EVQNWALNDHGHHRH 495
 RESULT 6
 ADR14587
 ID ADR14587 standard; protein: 495 AA.
 AC ADR14587;
 XX 21-OCT-2004 (first entry)
 DE Human NF-kappaB pathway-associated protein SeqID588.
 XX
 KM NF-kappaB pathway; antiinflammatory; cytosolic; hepatotropic; virucide;
 KM antiatheritic; antirheumatic; gastrointestinal-Gen; antiasthmatic;
 KM antiatherosclerotic; immunomodulator; cerebroprotective; vasotropic;
 KM immunosuppressive; vunerary; gene therapy; immune disorder;
 KM inflammatory disorder; NF-kappaB regulation; cancer; aberrant apoptosis;
 KM hepatic disorder; Hodgkin's lymphoma; haematopoietic tumour;
 KM hyper-IGM syndrome; hypohidrotic ectodermal dysplasia;
 KM X-linked andidrotic ectodermal dysplasia; immunodeficiency;
 KM viral infection; HIV-1; HTLV-1; hepatitis B; hepatitis C; EBV; influenza;
 KM viral replication; host cell survival; evasion of immune response;
 KM rheumatoid arthritis; inflammatory bowel disease; colitis; asthma;
 KM atherosclerosis; cachexia; euthyroid sick syndrome; stroke; EAE;
 KM autoimmune disorder; hyper immune activity;
 KM aberrant acute phase response; hypercongenital condition; birth defect;
 KM necrotic lesion; wound; organ transplant rejection;
 KM aberrant signal transduction; proliferating disorder; cancer;
 KM HIV propagation; human.
 XX
 OS Homo sapiens.
 OS
 XX
 PN WO2004065577-A2.
 PD 05-AUG-2004.
 XX
 PF 13-JAN-2004; 2004WO-US000798.
 XX
 PR 14-JAN-2003; 2003US-0440068P.
 PR 12-MAY-2003; 2003US-0469757P.
 XX
 PA (BRIM) BRISTOL-MYERS SQUIBB CO.
 XX
 PI Nadler SG, Neubauer MG, Feder JN, Carman J;

XX WPI; 2004-562168/54.
DR N-PSDB; ADRL4586.
XX
PT New isolated polynucleotides and polypeptides associated with NF-kappa
PT pathway, useful for diagnosing, treating, or preventing disorders or
PT diseases associated with NF-kappa pathway.
XX
PS Claim 6; SEQ ID NO 588; 237bp; English.
XX
XX This invention relates to the novel association of protein sequences (and
CC the genes which encode them) to the NF-kappa pathway. The invention may
CC be useful for the production of compounds with an antiinflammatory,
CC cytostatic, hepatotropic, virucide, antiarthritic, antirheumatic,
CC gastrointestinal-Gen, antiaesthetic, antiarteriosclerotic,
CC immunomodulator, cerebroprotective, vasotropic, immunosuppressive or
CC vulnerary activity or for gene therapy. The proteins and nucleotides are
CC useful for diagnosing, preventing, treating, or ameliorating conditions
CC or diseases associated with the NF-kappa pathway. The condition is an
CC immune disorder, an inflammatory disorder, cancer, aberrant apoptosis,
CC related to aberrant NF-kappa regulation, cancer, aberrant apoptosis,
CC hepatic disorders, Hodgkin's lymphomas, haematopoietic tumours, hyper-IgM
CC syndrome, hypohidrotic ectodermal dysplasia, X-linked anhidrotic
CC ectodermal dysplasia, immunodeficiency, viral infections, HIV-1, HTLV-1,
CC hepatitis B, hepatitis C, EBV, influenza, viral replication, host cell
CC survival, evasion of immune responses, rheumatoid arthritis, inflammatory
CC bowel disease, colitis, asthma, atherosclerosis, cachexia, euthyroid sick
CC syndrome, stroke, EAE, autoimmune disorders, disorders related to hyper
CC immune activity, disorders related to aberrant acute phase responses,
CC hydropregnant conditions, birth defects, necrotic lesions, wounds,
CC organ transplant rejection, conditions related to organ transplant
CC rejection, disorders related to aberrant signal transduction,
CC proliferating disorders, cancers and HIV propagation in cells infected
CC with other viruses. The present sequence is that of a human protein which
CC is subject to the novel association with the NF-kappa pathway of the
CC invention. Note: This sequence does not appear in the specification but
CC was obtained by the indexer from Genbank.
XX
XX
SQ Sequence 495 AA;
Query Match 92.4%; Score 2621; DB 8; Length 495;
Best Local Similarity 100.0%; Pred. No. 2,2e-260;
Matches 495; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 42 MRSPPVRLANDGESESTDRPPLFGAPRAEAPVCCSARVYLAALAFPGFIVYALRVNL 101
DB 1 MRSPPVRLANDGESESTDRPPLFGAPRAEAPVCCSARVYLAALAFPGFIVYALRVNL 60
QY 102 SVALVMDVDSNTLTEDNRTSKACPEHSAPIKVHNNGKKYQWMAEQGMLGSFFGYI 161
DB 61 SVALVMDVDSNTLTEDNRTSKACPEHSAPIKVHNNGKKYQWMAEQGMLGSFFGYI 120
QY 162 ITQIPGVYVASKIGKMLGFGILGTAVLTLFTPIADLVGLVPLIVRALGLEGVTFP 221
DB 121 ITQIPGVYVASKIGKMLGFGILGTAVLTLFTPIADLVGLVPLIVRALGLEGVTFP 180
QY 222 AMHAMSSWAPPLERSKILSTSYGAQGLTVISPLSGITTYVMWVYVFFFGTIGFW 281
DB 181 AMHAMSSWAPPLERSKILSTSYGAQGLTVISPLSGITTYVMWVYVFFFGTIGFW 240
QY 282 FLMIWIVSDTPQGHKRIHSEKEYIISLNNOSSQSVWVILKSLPMAIVVAHFS 341
DB 241 FLMIWIVSDTPQGHKRIHSEKEYIISLNNOSSQSVWVILKSLPMAIVVAHFS 300
QY 342 YNMTFYTLTLPTVMKEILRFNVQENGFLSSLPYLGSWLCMILSGQADNWRKMWEST 401
DB 301 YNMTFYTLTLPTVMKEILRFNVQENGFLSSLPYLGSWLCMILSGQADNWRKMWEST 360
QY 402 LCVARIFSLGIMGPVAVFLVAGFTICDYSIAVAFITSTLGGFCSSGSGINHLDIAP 461
DB 361 LCVARIFSLGIMGPVAVFLVAGFTICDYSIAVAFITSTLGGFCSSGSGINHLDIAP 420
QY 462 YAGILLGITNTFAITPGWGVPIAKSLTPDNTVGEWQVFYIAAIVFGAIFFTLFAKG 521

DB 421 YAGILLGITNTFAITPGWGVPIAKSLTPDNTVGEWQVFYIAAIVFGAIFFTLFAKG 480
QY 522 EYQWALNDHGHGRH 536
DB 481 EYQWALNDHGHGRH 495
RESULT 7
ID ADP25217 standard; protein; 495 AA.
XX
XX ADP25217;
AC
XX
XX 18-NOV-2004 (first entry)
DT
XX
DE PRO polypeptide SEQ ID NO:2395.
XX
XX PRO; antiinflammatory; antiarthritic; antirheumatic; immunosuppressive;
KW osteopathic; antidiabetic; dermatological; antipsoriatic; antiallergic;
KW antiaesthetic; hepatotropic; respiratory; gene therapy; immune system.
XX
XX Unidentified.
OS
XX
XX WC2004041170-A2.
PN
XX
XX 21-MAY-2004.
PD
XX
XX 30-OCT-2003; 2003WC-US034312.
PF
XX
XX 01-NOV-2002; 2002US-0423394P.
PR
XX
PA (GENTH) GENENTECH INC.
XX
XX Clark H, Schoenfeld J, Van Lookeren M, Williams PM, Wood WI;
PI Wu TD;
PI
XX
XX WPI; 2004-419628/39.
DR N-PSDB; ADP25216.
XX
XX
XX New PRO polypeptides and polynucleotides, useful for treating e.g.
PT erythematous, rheumatoid arthritis, diabetes mellitus, immune-mediated
PT renal disease, or demyelinating diseases of the central or peripheral
PT nervous system.
XX
PS Claim 7; SEQ ID NO 2395; 2940bp; English.
XX
XX The invention relates to a novel isolated nucleic acid and the PRO
CC polypeptide encoded by it. A protein of the invention has
CC antiinflammatory, antiarthritic, antirheumatic, immunosuppressive,
CC osteopathic, antidiabetic, dermatological, antipsoriatic, antiallergic,
CC antiaesthetic, hepatotropic, and respiratory activity. A polynucleotide
CC of the invention may have a use in gene therapy. The PRO polypeptide, its
CC agonist, antagonist, or antibody that specifically binds to the
CC polypeptide is useful for treating an immune related disorder such as
CC systemic lupus erythematous, rheumatoid arthritis, osteoarthritis,
CC juvenile chronic arthritis, a spondyloarthropathy, systemic sclerosis, an
CC idiopathic inflammatory myopathy, Sjogren's syndrome, systemic
CC vasculitis, sarcoidosis, autoimmune haemolytic anaemia, autoimmune
CC thrombocytopenia, thyroiditis, diabetes mellitus, immune-mediated renal
CC disease, a demyelinating disease of the central or peripheral nervous
CC system, idiopathic demyelinating polyneuropathy, Guillain-Barre syndrome,
CC a chronic inflammatory demyelinating polyneuropathy, a hepatobiliary
CC disease, infectious or autoimmune chronic active hepatitis, primary
CC biliary cirrhosis, granulomatous hepatitis, sclerosing cholangitis,
CC inflammatory bowel disease, gluten-sensitive enteropathy, Whipple's
CC disease, an autoimmune or immune-mediated skin disease, a bullous skin
CC disease, erythema multiforme, contact dermatitis, psoriasis, an allergic
CC disease, asthma, allergic rhinitis, atopic dermatitis, food
CC hypersensitivity, urticaria, an immunologic disease of the lung,
CC eosinophilic pneumonia, idiopathic pulmonary fibrosis, hypersensitivity
CC pneumonitis, a transplantation associated disease, graft rejection or
CC graft-versus-host disease. The present sequence represents a PRO protein

CC of the invention.
XX
SQ Sequence 495 AA;

Query Match 92.4%; Score 2621; DB 8; Length 495;
Best Local Similarity 100.0%; Pred. No. 2.2e-260;
Matches 495; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 42 MRSFVRDLARNDDGSESTDRTPLLPGAPRAEAPVCCSARNYLAIALAFGFPIYALRVNL 101
DB 1 MRSFVRDLARNDDGSESTDRTPLLPGAPRAEAPVCCSARNYLAIALAFGFPIYALRVNL 60
QY 102 SVALVDVDSNTLTEDNRTSKACEHSAPIKVHHNQGKKYQMDAETQGMILGSFFGYI 161
DB 61 SVALVDVDSNTLTEDNRTSKACEHSAPIKVHHNQGKKYQMDAETQGMILGSFFGYI 120
QY 162 ITQIPGGYVASKIGKMLGFGILGTAVLTLFPIADLGVGPLIVRALEGLGEGVTFP 221
DB 121 ITQIPGGYVASKIGKMLGFGILGTAVLTLFPIADLGVGPLIVRALEGLGEGVTFP 180
QY 222 AMHAMSSWAPPLERSKULSISYAGQOLGVISLPLSGIICVYNNMTYVFFPGTIGIFW 281
DB 181 AMHAMSSWAPPLERSKULSISYAGQOLGVISLPLSGIICVYNNMTYVFFPGTIGIFW 240
QY 282 FLMIWLVSDTPQKHKRISHYEKEYILSLRNQLSQKSPWVPILKSLPLMAIVVAHFS 341
DB 241 FLMIWLVSDTPQKHKRISHYEKEYILSLRNQLSQKSPWVPILKSLPLMAIVVAHFS 300
QY 342 YNNTFYTLTLPTPYMKEILRFNVQENGFLSSLPYLGSWLCMILSGAADNLRKNNFST 401
DB 301 YNNTFYTLTLPTPYMKEILRFNVQENGFLSSLPYLGSWLCMILSGAADNLRKNNFST 360
QY 402 LCVRRIFSLIGMIGPAFLVAAAGTIGCDYSLAAVFLTISTTLGFCSSGSGFINHLDIAPS 461
DB 361 LCVRRIFSLIGMIGPAFLVAAAGTIGCDYSLAAVFLTISTTLGFCSSGSGFINHLDIAPS 420
QY 462 YAGILGITTTFATIPGMVGPVIAKSLTPDNTVGEMQTVFYIAAINVFAGIFFTLPAKG 521
DB 421 YAGILGITTTFATIPGMVGPVIAKSLTPDNTVGEMQTVFYIAAINVFAGIFFTLPAKG 480
QY 522 EVQNMALNDHGHHRH 536
DB 481 EVQNMALNDHGHHRH 495
```

RESULT 8
AAM38959
ID AAM38959 standard; protein; 495 AA.
XX
AC AAM38959;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 2104.
XX
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; haemoblastic; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic; chemokinetic; thrombolytic; drug screening; arthritis; inflammation; leukaemia.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US034263.
XX
PR 23-DEC-1999; 99US-00471275.
PR 21-JAN-2000; 2000US-00488725.
PR 25-APR-2000; 2000US-0052317.

```
PR 20-JUN-2000; 2000US-00598042.  
PR 19-JUL-2000; 2000US-00620312.  
PR 03-AUG-2000; 2000US-00653450.  
PR 14-SEP-2000; 2000US-00662191.  
PR 19-OCT-2000; 2000US-00693036.  
PR 29-NOV-2000; 2000US-00727344.  
XX  
XX (HYSE-) HYSEQ INC.  
PA Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D,  
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AU, Yang Y, Zhang J, Zhao QA,  
PI Zhou P, Goodrich R, Drmanac RT;  
XX WPI: 2001-442253/47.  
DR N-PSDB; AA158115.  
XX  
XX Novel nucleic acids and polypeptides, useful for treating disorders such  
PT as central nervous system injuries.  
XX  
XX Example 3; SEQ ID NO 2104; 10078pp; English.  
PS  
XX The invention relates to human nucleic acids (AA157798-AA161365) and the  
CC encoded polypeptides (AAM38642-AA42213) with nootropic,  
CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
CC in gene therapy. A composition containing a polypeptide or polynucleotide  
CC of the invention may be used to treat diseases of the peripheral nervous  
CC system, such as peripheral nervous injuries, peripheral neuropathy and  
CC localised neuropathies and central nervous system diseases, such as  
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
CC utilisation of the activities such as: Immune system suppression,  
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
CC assays for receptor activity, arthritis and inflammation, leukaemia and  
CC C.N.S disorders. Note: The sequence data for this patent did not form  
XX part of the printed specification  
SQ  
Sequence 495 AA;  
QY 42 MRSFVRDLARNDDGSESTDRTPLLPGAPRAEAPVCCSARNYLAIALAFGFPIYALRVNL 101  
DB 1 MRSFVRDLARNDDGSESTDRTPLLPGAPRAEAPVCCSARNYLAIALAFGFPIYALRVNL 60  
QY 102 SVALVDVDSNTLTEDNRTSKACEHSAPIKVHHNQGKKYQMDAETQGMILGSFFGYI 161  
DB 61 SVALVDVDSNTLTEDNRTSKACEHSAPIKVHHNQGKKYQMDAETQGMILGSFFGYI 120  
QY 162 ITQIPGGYVASKIGKMLGFGILGTAVLTLFPIADLGVGPLIVRALEGLGEGVTFP 221  
DB 121 ITQIPGGYVASKIGKMLGFGILGTAVLTLFPIADLGVGPLIVRALEGLGEGVTFP 180  
QY 222 AMHAMSSWAPPLERSKULSISYAGQOLGVISLPLSGIICVYNNMTYVFFPGTIGIFW 281  
DB 181 AMHAMSSWAPPLERSKULSISYAGQOLGVISLPLSGIICVYNNMTYVFFPGTIGIFW 240  
QY 282 FLMIWLVSDTPQKHKRISHYEKEYILSLRNQLSQKSPWVPILKSLPLMAIVVAHFS 341  
DB 241 FLMIWLVSDTPQKHKRISHYEKEYILSLRNQLSQKSPWVPILKSLPLMAIVVAHFS 300  
QY 342 YNNTFYTLTLPTPYMKEILRFNVQENGFLSSLPYLGSWLCMILSGAADNLRKNNFST 401  
DB 301 YNNTFYTLTLPTPYMKEILRFNVQENGFLSSLPYLGSWLCMILSGAADNLRKNNFST 360  
QY 402 LCVRRIFSLIGMIGPAFLVAAAGTIGCDYSLAAVFLTISTTLGFCSSGSGFINHLDIAPS 461  
DB 361 LCVRRIFSLIGMIGPAFLVAAAGTIGCDYSLAAVFLTISTTLGFCSSGSGFINHLDIAPS 420  
QY 462 YAGILGITTTFATIPGMVGPVIAKSLTPDNTVGEMQTVFYIAAINVFAGIFFTLPAKG 521  
DB 421 YAGILGITTTFATIPGMVGPVIAKSLTPDNTVGEMQTVFYIAAINVFAGIFFTLPAKG 480
```

Query Match 92.2%; Score 2614; DB 4; Length 495;
Best Local Similarity 99.8%; Pred. No. 1.2e-259;
Matches 494; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Db	421	YAGILLGTTNTFATLPGMGPVIAKSLTPDNTVGEMQVFYIIAAINVEGAIFITLEFAKG
Oy	522	EYQNMMALNDHGHRR 536
Db	481	EYQNMMALNDHGHRR 495
RESULT 9		
ID	AAY45088	standard; protein; 495 AA.
XX	AAY45088;	
XX	31-MAY-2000	(first entry)
XX		
DE		Sheep GBS toxin receptor (SP55).
XX		
KW		Sheep GBS toxin receptor; group B beta-haemolytic streptococci; SP55;
KW		pathological vasculaturation; cancer metastases; angiogenesis; sheep;
KW		neovascularisation; reperfusion injury; scarring; keloid;
KW		chronic inflammation; rheumatoid arthritis; psoriasis; neural injury;
KW		endothelial cell proliferation; antibacterial; anticancer;
KW		anti-angiogenic; anti-inflammatory; anti-arthritis; anti-psoriatic.
XX		
OS	Ovis sp.	
XX		
FH	Key	Location/Qualifiers
FT	Peptide	7..22 /label= p55a /note= "Immunogenic peptide"
FT		/note= "Immunogenic peptide"
FT	Peptide	8..35 /label= p56a /note= "Immunogenic peptide"
FT		/note= "Immunogenic peptide"
FT	Modified-site	11..14 /label= CK2.phospho.site /note= "Predicted phosphorylation site"
FT		/note= "Predicted phosphorylation site"
FT	Modified-site	17..19 /label= PKC.phospho.site /note= "Predicted phosphorylation site"
FT		/note= "Predicted phosphorylation site"
FT	Modified-site	37..39 /label= PKC.phospho.site /note= "Predicted phosphorylation site"
FT		/note= "Predicted phosphorylation site"
FT	Domain	42..58 /note= " Predicted transmembrane domain"
FT		/note= " Predicted transmembrane domain"
FT	Modified-site	55..57 /label= PKC.phospho.site /note= "Predicted phosphorylation site"
FT		/note= "Predicted phosphorylation site"
FT	Modified-site	59..62 /note= "Asn is N-glycosylated"
FT		/note= "Asn is N-glycosylated"
FT	Peptide	71..84 /label= p57a /note= "Immunogenic peptide"
FT		/note= "Immunogenic peptide"
FT	Modified-site	71..74 /note= "Asn is N-glycosylated"
FT		/note= "Asn is N-glycosylated"
FT	Modified-site	73..76 /label= CK2.phospho.site /note= "Predicted phosphorylation site"
FT		/note= "Predicted phosphorylation site"
FT	Modified-site	73..75 /label= PKC.phospho.site /note= "Predicted phosphorylation site"
FT		/note= "Predicted phosphorylation site"
FT	Modified-site	77..80 /note= "Asn is N-glycosylated"
FT		/note= "Asn is N-glycosylated"
FT	Modified-site	79..82 /label= CK2.phospho.site /note= "Predicted phosphorylation site"
FT		/note= "Predicted phosphorylation site"
FT	Modified-site	95..98 /note= "Asn is N-glycosylated"
FT		/note= "Asn is N-glycosylated"
FT	Modified-site	97..100 /note= "Putative amidation site"
FT		/note= "Putative amidation site"
FT	Modified-site	97..99 /label= PKC.phospho.site /note= "Predicted phosphorylation site"
FT		/note= "Predicted phosphorylation site"
FT	Modified-site	126..131

FT		/note= "Predicted myristylation site"
FT	Domain	135. .157
FT	/note= "Outer boundary of transmembrane domain"	137. .153
FT	Domain	137. .153
FT	Modified-site	/note= "Inner boundary of transmembrane domain"
FT		142. .147
FT	Modified-site	/note= "Predicted myristylation site"
FT	Modified-site	162. .167
FT	Modified-site	/note= "Predicted myristylation site"
FT	Modified-site	172. .177
FT	Modified-site	/note= "Predicted myristylation site"
FT	Modified-site	205. .210
FT	Modified-site	/note= "Predicted myristylation site"
FT	Modified-site	209. .214
FT	Modified-site	/note= "Predicted myristylation site"
FT	Modified-site	225. .228
FT	Modified-site	/note= "Aen is N-glycosylated"
FT	Domain	226. .252
FT		/note= "Outer boundary of transmembrane domain"
FT	Domain	232. .248
FT	Modified-site	/note= "Inner boundary of transmembrane domain"
FT	Modified-site	254. .256
FT	Modified-site	/label= PKC phospho site
FT	Modified-site	/note= "Predicted phosphorylation site"
FT	Modified-site	259. .262
FT	Modified-site	/label= CK2 phospho site
FT	Modified-site	/note= "Predicted phosphorylation site"
FT	Modified-site	269. .271
FT	Modified-site	/label= PKC phospho site
FT	Modified-site	/note= "Predicted phosphorylation site"
FT	Modified-site	276. .278
FT	Modified-site	/label= PKC phospho site
FT	Modified-site	/note= "Predicted phosphorylation site"
FT	Modified-site	302. .305
FT	Modified-site	/note= "Aen is N-glycosylated"
FT	Domain	328. .345
FT		/note= "Outer boundary of transmembrane domain"
FT	Domain	328. .344
FT	Modified-site	/note= "Inner boundary of transmembrane domain"
FT	Modified-site	337. .342
FT	Modified-site	/note= "Predicted myristylation site"
FT	Modified-site	357. .360
FT	Modified-site	/note= "Aen is N-glycosylated"
FT	Domain	365. .389
FT		/note= "Outer boundary of transmembrane domain"
FT	Domain	369. .385
FT	Modified-site	/note= "Inner boundary of transmembrane domain"
FT	Modified-site	386. .391
FT	Modified-site	/note= "Predicted myristylation site"
FT	Domain	390. .407
FT		/note= "Outer boundary of transmembrane domain"
FT	Domain	390. .406
FT	Modified-site	/note= "Inner boundary of transmembrane domain"
FT	Modified-site	403. .408
FT	Modified-site	/note= "Predicted myristylation site"
FT	Modified-site	423. .428
FT	Modified-site	/note= "Predicted myristylation site"
FT	Modified-site	427. .432
FT	Modified-site	/note= "Predicted myristylation site"
FT	Modified-site	452. .455
FT	Modified-site	/label= CK2 phospho site
FT	Domain	/note= "Predicted phosphorylation site"
FT	Domain	456. .479
FT	Domain	/note= "Outer boundary of transmembrane domain"
FT	Domain	458. .474
FT		/note= "Inner boundary of transmembrane domain"
XX	WO200005375-A1.	
XX	03-FEB-2000.	
XX	22-JUL-1999;	99WO-US016676.

PR 22-JUN-1998; 98US-0093843P.
 XX (UYVA-) UNIV VANDERBILT.
 PA
 XX
 PI Hellerqvist CG, Fu C;
 XX
 XX WPI; 2000-205377/18.
 DR N-PSDB; AAZ50876.
 XX
 XX
 PT New polynucleotide encoding mammalian receptor for streptococcus toxin,
 PT useful for diagnosis and treatment of, e.g. pneumonia in neonates.
 XX
 XX
 PS Claim 10; Page 22; 109pp; English.
 XX
 CC The present sequence is sheep GBS (group B beta-haemolytic streptococci)
 CC toxin receptor (SP55). Sheep GBS toxin receptor is an integral protein
 CC with seven transmembrane domains. Expression vectors comprising the
 CC coding region can be transformed into host cells to express GBS toxin
 CC receptor and its fragments. Detecting the receptor in tissues is used to
 CC diagnose pathological vascularisation, e.g. for detecting cancer
 CC metastases. GBS toxin receptors are useful for treating conditions
 CC associated with pathological angiogenesis or neovascularisation
 CC (specifically cancer, reperfusion injury, scarring during wound healing,
 CC keloids, chronic inflammation (rheumatoid arthritis or psoriasis) or
 CC neural injury), and to raise specific antibodies used for treating early
 CC onset disease. Inhibitors of this receptor are useful for treating
 CC pathological or hypoxia-induced endothelial cell proliferation and
 CC migration
 CC
 SQ Sequence 495 AA;
 Query Match 82.1%; Score 2329; DB 3; Length 495;
 Best Local Similarity 86.3%; Pred. No. 2.6e-230;
 Matches 427; Conservative 39; Mismatches 22; Indels 0; Gaps 0;
 QY 42 MRSVPRDLARNDGEESTDRTPPLPGAPRAEAPVCCSARNYLAIAFFGFIVYALRVNL 101
 DB 1 MKSPVSLDAPSDGEGSDRTPLORAPRAEPAPVCCSARNYLAIAFFGFIVYALRVNL 60
 QY 102 SVALVDVNDSTLTLEDNRTSKACEPSHAPIRVHNQGTQOMAEITGWLGSFFGYXI 161
 DB 61 SVALVDVNDSTLTAKDNRTSYECASHAPIRVHNQGTQOMAEITGWLGSFFGYXI 120
 QY 162 ITQIPGGVASKIGKMLGFGIIGTAVLTFTPLAADLGVPLIVLRALGEGVTEP 221
 DB 121 ITQIPGGVASKIGKMLGFGIIGTAVLTFTPLAADLGVPLIVLRALGEGVTEP 180
 QY 222 AMHAMSSWAPPLERSKLLSISVAGAOIGTVISLPLSGIICYNNMTVFFYFPGTIGIFW 281
 DB 181 AMHAMSSWAPPLERSKLLSISVAGAOIGTVISLPLSGIICYNNMTVFFYFPGTIGIFW 240
 QY 282 FLMTMLVSDTPQGHKASHKEYIYSSLENOJSSQGSVWVPLKSLPMAIVVAHS 341
 DB 241 FLMTMLVSDTPQGHKASHKEYIYSSLENOJSSQGSVWVPLKSLPMAIVVAHS 300
 QY 342 YNMTFTYLLTLPYMKELIRENVOENGFLSLPYLGSMLCMILSGOANDLRKMNFT 401
 DB 301 YNMTFTYLLTLPYMKELIRENVOENGFLSLPYLGSMLCMILSGOANDLRKMNFT 360
 QY 402 LCVARISLIGMIGPAVFLVAGFICGDYSIAVAFITISTLGGFCSGSGFINHDIAPS 461
 DB 361 LCVARISLIGMIGPAVFLVAGFICGDYSIAVAFITISTLGGFCSGSGFINHDIAPS 420
 QY 462 YAGILGITYTFAITPBGVGVIAKSLTPDNTVGEMQVFYIAAIVFGAIFFTLPKAG 521
 DB 421 YAGILGITYTFAITPBGVGVIAKSLTPDNTVGEMQVFYIAAIVFGAIFFTLPKAG 480
 QY 522 EVQNMALNDHGHGRH 536
 DB 481 EVQNMALNDHGHGRH 495

AAE06519
 ID AAE06519 standard; protein; 495 AA.
 XX
 AC AAE06519;
 XX
 DT 16-OCT-2001 (first entry)
 XX
 DE Sheep group B beta-haemolytic Streptococci toxin receptor (SP55) protein.
 XX
 KW Sheep, group B beta-haemolytic Streptococci toxin receptor; GBS; SP55;
 KW cytosolic; vulnerrary; antiatherosclerotic; osteopathic; vasotropic;
 KW prevention; attenuation; pathoangiogenic condition; cancer; scar;
 KW wound healing; gliosis; nerve injury; chronic wound; reperfusion injury;
 KW keloid; rheumatoid arthritis; atherosclerosis; osteoarthritis; psoriasis;
 KW vaccine.
 KW
 OS Ovis sp.
 XX
 FH Key
 FT Region
 FT Location/Qualifiers
 FT 8..22
 FT /label= p55a_immunogenic_peptide
 FT /note= "Fragment of extracellular domain of GBS toxin
 FT receptor"
 FT 9..35
 FT /label= p56a_immunogenic_peptide
 FT 14..19
 FT /note= "Region of high hydrophilicity"
 FT 25..30
 FT /note= "Region of high hydrophilicity"
 FT 71..84
 FT /label= p57a_immunogenic_peptide
 FT /note= "Fragment of intracellular domain of GBS toxin
 FT receptor"
 FT 75..80
 FT /note= "Region of high hydrophilicity"
 FT Region
 FT
 FT WO200156598-A2.
 XX
 PN 09-AUG-2001.
 XX
 PD 02-FEB-2001; 2001WO-US003662.
 XX
 PF 02-FEB-2000; 2000US-0179870P.
 XX
 PR (UYVA-) UNIV VANDERBILT.
 XX
 PA
 XX
 PI Hellerqvist CG;
 XX
 DR WPI; 2001-488844/53.
 XX
 DR N-PSDB; AAD10326.
 XX
 PT Preventing or attenuating pathoangiogenic conditions e.g. cancer, chronic
 PT wounds, osteoarthritis, keloids and psoriasis in a mammal, by
 PT administering group B beta-haemolytic Streptococci toxin receptor or its
 PT fragment.
 XX
 PS Claim 8; Page 50-52; 52pp; English.
 XX
 CC The present sequence is group B beta-haemolytic Streptococci (GBS) toxin
 CC receptor protein, SP55 from sheep. The present invention relates to a
 CC method for preventing or attenuating a patho-angiogenic condition in a
 CC mammal which comprises administering to the mammal one or more GBS toxin
 CC receptors or their immunogenic fragments to induce or maintain an immune
 CC response to one of GBS toxin receptors. The method is useful for
 CC preventing or ameliorating pathoangiogenic conditions such as cancer,
 CC scarring during wound healing, gliosis during repair of nerve injury,
 CC chronic wounds, keloids, reperfusion injury, rheumatoid arthritis,
 CC atherosclerosis, osteoarthritis and psoriasis in mammals. The proteins of
 CC the invention are also used as vaccines
 CC
 SQ Sequence 495 AA;
 Query Match 82.1%; Score 2329; DB 4; Length 495;

AC AAG65238;
XX 20-NOV-2001 (first entry)
XX
DE Human sodium dependent phosphate co-transfer protein 35.
XX
XX Human; sodium dependent phosphate co-transfer protein 35;
KM hypophosphaturia; hypercalcaemia; hypophosphataemic rickets; nephritis;
KM gene therapy.
XX
OS Homo sapiens.
PN CN1298882-A.
XX
XX 13-JUN-2001.
XX
XX 06-DEC-1999; 99CN-00124217.
XX
XX 06-DEC-1999; 99CN-00124217.
XX
XX 06-DEC-1999; 99CN-00124217.
XX
XX (SHEN-) SHENGYUAN GENE DEV CO LTD SHANGHAI.
XX
XX Mao Y, Xie Y;
XX
XX WPI; 2001-503367/56.
DR N-PSDB; AAH79234.
XX
XX Human Na-dependent phosphate cotransporter 35 and its coding sequence.
PT
XX
XX Claim 1; Page 21 (Disclosure); 28pp; Chinese.
XX
XX The present invention provides the protein and coding sequences of human
CC sodium dependent phosphate co-transfer protein 35. The sequences can be
CC used in the treatment of hypophosphaturia, hypercalcaemia,
CC hypophosphataemic rickets and nephritis. The present sequence is the
CC protein of the invention
XX
XX
SQ Sequence 314 AA;

Query Match 59.2%; Score 1679; DB 4; Length 314;
Best Local Similarity 100.0%; Pred. No. 1.2e-163; Mismatches 0; Gaps 0;
Matches 314; Conservative 0; Indels 0; Gaps 0;

QY 223 MHAMMSMAPLEKSKLISISYAGOLGVISLPLSGIICYMMWTVYFVFFGTIGTFWF 282
DB 1 MHAMMSMAPLEKSKLISISYAGOLGVISLPLSGIICYMMWTVYFVFFGTIGTFWF 60
QY 283 LWMWLVSDTPQKHKRIISHEKEYILSLRNQSSOKSVPMVPIKSLPLMAIVVAHFSY 342
DB 61 LWMWLVSDTPQKHKRIISHEKEYILSLRNQSSOKSVPMVPIKSLPLMAIVVAHFSY 120
QY 343 NMFFYTLTLPTMKELIRNVQENGLSLPLTSGMLCMILSGOANDNRKMNSTL 402
DB 121 NMFFYTLTLPTMKELIRNVQENGLSLPLTSGMLCMILSGOANDNRKMNSTL 180
QY 403 CVRRIFSLIGMIGAVFLVAAFGICDYSALAVFLTISTTLGFCSSGFSINHDIPSY 462
DB 181 CVRRIFSLIGMIGAVFLVAAFGICDYSALAVFLTISTTLGFCSSGFSINHDIPSY 240
QY 463 AGILGTTNTFATIPGAVGVIAKSLTPDNTVGEMQTFVYIAAINVGAIFFTLPAKE 522
DB 241 AGILGTTNTFATIPGAVGVIAKSLTPDNTVGEMQTFVYIAAINVGAIFFTLPAKE 300
QY 523 VQNMALNDHHRH 536
DB 301 VQNMALNDHHRH 314

RESULT 13
ID AAM40745 standard; protein; 309 AA.
XX
XX AAM40745;
AC

XX 22-OCT-2001 (first entry)
DT
XX
XX Human polypeptide SEQ ID NO 5676.
DE
XX
XX Human; noctropic; immunosuppressant; cytostatic; gene therapy; cancer;
KM peripheral nervous system; neuropathy; central nervous system; CNS;
KM Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KM amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KM chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KM leukaemia.
XX
XX Homo sapiens.
OS
XX
XX WO200153312-A1.
PN
XX
XX 26-JUL-2001.
PD
XX
XX 26-DEC-2000; 2000WO-US034263.
XX
XX 23-DEC-1999; 99US-00471275.
PR 21-JAN-2000; 2000US-00489725.
PR 25-APR-2000; 2000US-00552317.
PR 20-JUN-2000; 2000US-00598042.
PR 19-JUL-2000; 2000US-00620312.
PR 03-AUG-2000; 2000US-00653450.
PR 14-SEP-2000; 2000US-00662191.
PR 19-OCT-2000; 2000US-00693036.
PR 29-NOV-2000; 2000US-00727344.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI Zhou F, Goodrich R, Drmanac RT;
XX
XX WPI; 2001-442253/47.
DR N-PSDB; AAI59901.
DR
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders such
PT as central nervous system injuries.
XX
XX Example 2; SEQ ID NO 5676; 10078pp; English.
PS
XX
XX The invention relates to human nucleic acids (AA157798-AA161369) and the
CC encoded polypeptides (AAM38642-AAM42213) with noctropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening
CC assays for receptor activity, arthritis and inflammation, leukaemia and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification
XX
SQ Sequence 309 AA;

Query Match 56.9%; Score 1614; DB 4; Length 309;
Best Local Similarity 100.0%; Pred. No. 5.9e-157; Mismatches 0; Gaps 0;
Matches 304; Conservative 0; Indels 0; Gaps 0;

QY 8 PPRPVQPARPGFGLSGRSLICQVASTPAVGMSPVVDLARNDEESTDRTPLPGA 67
DB 6 PPRPVQPARPGFGLSGRSLICQVASTPAVGMSPVVDLARNDEESTDRTPLPGA 65
QY 68 PRABAPVCCSARYNLAIIAFAFGFFIYALRVNLVALVUMVDSNTTLENNRISKACPEH 127
DB 66 PRABAPVCCSARYNLAIIAFAFGFFIYALRVNLVALVUMVDSNTTLENNRISKACPEH 125

QY 128 SAPIKHHNQTGKKYQWDATQGWILGSPFYGYITITQPGYVASKIGKMLGFGILGT 187
 CC |||||
 CC 126 SAPIKHHNQTGKKYQWDATQGWILGSPFYGYITITQPGYVASKIGKMLGFGILGT 185
 CC |||||
 QY 188 AVITLFTPIAADGVGPLYLRLALEGEGVTPFAMHAMSSMAPPLERSKLISYAGA 247
 CC |||||
 Db 186 AVITLFTPIAADGVGPLYLRLALEGEGVTPFAMHAMSSMAPPLERSKLISYAGA 245
 CC |||||
 QY 248 QLGTVISLPLSGIICYYMMNTYVFFPGTIGIFWFLIMLVSDTPQKHKISHYEKEYI 307
 CC |||||
 Db 246 QLGTVISLPLSGIICYYMMNTYVFFPGTIGIFWFLIMLVSDTPQKHKISHYEKEYI 305
 CC |||||
 QY 308 LSSL 311
 CC |||||
 Db 306 LSSL 309
 CC |||||
 RESULT 14
 ABP41345
 ID ABP41345 standard; protein; 284 AA.
 XX
 AC ABP41345;
 XX
 DT 22-AUG-2002 (first entry)
 XX
 DE Human ovarian antigen HMWCG79, SEQ ID NO:2477.
 XX
 XX Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
 KM ovarian cancer; breast cancer; tumour; reproductive system disorder;
 KM infertility; pregnancy disorders; anovulation; polycystic ovary syndrome;
 KM PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;
 KM inflammatory condition; immune disorder; blood disorder;
 KM cardiovascular disorder; respiratory disorder; neurological disorder;
 KM gastrointestinal disorder; urinary system disorder; drug screening;
 KM gene therapy; chromosome mapping; forensic analysis;
 KM antibody preparation; cytostatic; immunomodulatory; neuroprotective;
 KM antiinflammatory; gynaecological; reproductive.
 KM
 OS Homo sapiens.
 XX
 PN WO200200677-A1.
 XX
 PD 03-JAN-2002.
 XX
 PF 07-JUN-2001; 2001WO-US018569.
 XX
 PR 07-JUN-2000; 2000US-0209467P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Birse CE, Rosen CA;
 XX
 DR WPI; 2002-147878/19.
 XX
 DR N-PSDB; AB054422.
 XX
 PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,
 PT useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian
 PT cancer), immune disorders, cardiovascular disorders and neurological
 PT diseases.
 XX
 PS Claim 11; SEQ ID NO 2477; 2922bp; English.
 XX
 CC The invention relates to 2175 novel human ovarian antigens (ABP41054-
 CC ABP43228) and to cDNAs encoding them (AB054131-AB056305), and also
 CC encompasses polypeptides 90% identical and polynucleotides 95% identical
 CC to the sequences of the invention. The invention additionally relates to
 CC recombinant vectors and host cells comprising human ovarian antigen
 CC polynucleotides, antibodies against human ovarian antigens, and the use
 CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
 CC treating, prognosing or preventing various ovary and/or breast-related
 CC disorders. Such conditions include ovarian cancer and breast cancer, and
 CC metastatic tumours of ovarian or breast origin, reproductive system

CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
 CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine
 CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
 CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
 CC vaginitis), immune disorders (e.g., congenital and acquired
 CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
 CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
 CC respiratory disorders, neurological disorders, gastrointestinal disorders
 CC and urinary system disorders. Ovarian antigen polypeptides and
 CC polynucleotides may also be used in screening for compounds which
 CC modulate ovarian antigen expression or activity. The polynucleotides may
 CC further be used for gene therapy, chromosome mapping, in the
 CC identification of individuals and in forensic analysis, and the
 CC polypeptides may be used as food additives or to prepare antibodies
 CC useful in disease diagnosis, drug targeting and phenotyping. The present
 CC sequence represents a human ovarian antigen of the invention. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences
 XX

SO Sequence 284 AA;

Query Match 53.6%; Score 1521; DB 5; Length 284;

Best Local Similarity 100.0%; Pred. No. 2e-147;

Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 253 ISLPLSGIICYYMMNTYVFFPGTIGIFWFLIMLVSDTPQKHKISHYEKEYIISLR 312

Db 1 ISLPLSGIICYYMMNTYVFFPGTIGIFWFLIMLVSDTPQKHKISHYEKEYIISLR 60

QY 313 NQSSQKSVBWDILKSLPMAIVAHFSYNTFTYLLTLPYMKELRFNVOENGFLS 372

Db 61 NQSSQKSVBWDILKSLPMAIVAHFSYNTFTYLLTLPYMKELRFNVOENGFLS 120

QY 373 SLPLSGMLCMISGQADNLRKKNPSTLCVRIRISLGMGPAPFLVAAGFTGCDYSL 432

Db 121 SLPLSGMLCMISGQADNLRKKNPSTLCVRIRISLGMGPAPFLVAAGFTGCDYSL 180

QY 433 AVAFPLTISTLTGGFCSSGFISNHLDIAPSYAGILGITNFPATIPGVGVIAKSLTPDN 492

Db 181 AVAFPLTISTLTGGFCSSGFISNHLDIAPSYAGILGITNFPATIPGVGVIAKSLTPDN 240

QY 493 TVGEMQTVFYIAAIVFGAIFFTLPFAKGEVQWALNDHGHRRH 536

Db 241 TVGEMQTVFYIAAIVFGAIFFTLPFAKGEVQWALNDHGHRRH 284

RESULT 15

ABB89975
 ID ABB89975 standard; protein; 272 AA.

XX
 AC ABB89975;

XX
 DT 24-MAY-2002 (first entry)

XX
 DE Human polypeptide SEQ ID NO 2351.

XX
 KM Cytostatic; immunosuppressive; nocrotropic; neuroprotective; antiviral;
 KM antiallergic; hepatotropic; antidiabetic; antiinflammatory; anticancer;
 KM vulnerrary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KM cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KM neurological disease; infection; human; secreted protein.

XX
 OS Homo sapiens.

XX
 PN WO200190304-A2.

XX
 PD 29-NOV-2001.

XX
 PF 18-MAY-2001; 2001WO-US016450.

XX
 PR 19-MAY-2000; 2000US-0205515P.

XX
 CC

PA (HUMA-) HUMAN GENOME SCI INC.

XX Birse CE, Rosen CA;

XX WPI; 2002-122018/16.

DR N-PSDB; ABL90384.

XX Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
PT prevention of neural, immune system, muscular, reproductive,
PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative
PT disorders.

PS Claim 11; SEQ ID NO 2351; 2081bp + Sequence Listing; English.

XX The invention relates to novel genes (ABL99449-ABL90853) and proteins
CC (AB89040-AB89044) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing
CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC infectious diseases such as viral, bacterial, fungal and parasitic
CC infections. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 272 AA;

Query Match 51.2%; Score 1451; DB 5; Length 272;

Best Local Similarity 99.6%; Pred. No. 3e-140; Mismatches 1; Indels 0; Gaps 0;

Matches 271; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 265 MNWTVYVFPGCTIGIFWFLMIWLVSDTPQKHRIKSHYEKEYIISLRLNQLSSQKSVPMV 324

DB 1 MNWTVYVFPGCTIGIFWFLMIWLVSDTPQKHRIKSHYEKEYIISLRLNQLSSQKSVPMV 60

QY 325 PILKSLPLMAIVAHFSSYNTFTYLLPTLPTMKEILRFNVQENGFLSSLPYLGSWLCMI 384

DB 61 PILKSLPLMAIVAHFSSYNTFTYLLPTLPTMKEILRFNVQENGFLSSLPYLGSWLCMI 120

QY 385 LSGQPADNLRKMNFTLCVRRIFSLIGMIGPAVFLVAAGFIGCDYSLAVAFITISTLG 444

DB 121 LSGQPADNLRKMNFTLCVRRIFSLIGMIGPAVFLVAAGFIGCDYSLAVAFITISTLG 180

QY 445 GFCSGSGSINHLDIAPSYAGILGINTFATIPGMNGPVIKSLTPDNTVGEMQTVFYIA 504

DB 181 GFCSGSGSINHLDIAPSYAGILGINTFATIPGMNGPVIKSLTPDNTVGEMQTVFYIA 240

QY 505 AAINVFGAIFFTLLFAKGEVQNMALNDHGHRRH 536

DB 241 AAINVFGAIFFTLLFAKGEVQNMALNDHGHRRH 272

Search completed: June 7, 2006, 05:37:46
Job time : 201 secs

THIS PAGE BLANK (USPTO)

November 2005

Published_Applications_Nucleic Acid and Published_Applications_Amino Acid database searches now generate two sets of results each. The Published_Applications databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Newly published applications will appear in the Published_Applications_New databases; older published applications make up the Published_Applications_Main databases.

Searches run against Nucleic Acid Published_Applications produce two sets of results, with the extensions **.rnpbm** (Published_Applications_NA_Main) and **.rnpbn** (Published_Applications_NA_New).

Searches run against Amino Acid Published_Applications produce two sets of results, with the extensions **.rapbm** (Published_Applications_AA_Main) and **.rapbn** (Published_Applications_AA_New).

THIS PAGE BLANK (USPTO)

GenCore version 5.1.9
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using SW model

Run on: June 7, 2006, 05:43:57 ; Search time 188 Seconds
(without alignments)
1320.655 Million cell updates/sec

Title: US-10-823-506-8

Perfect score: 2836

Sequence: 1 MAGAMTPPRPVQPARPGGF.....LFAKGEVQWALNDHGHRRH 536

Scoring table: BLOSUM62

Searched: 2097797 seqs, 463214858 residues

Total number of hits satisfying chosen parameters: 2097797

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications_AA_Main:*

1: /EMC_Celerra_SIDS3/prodata/2/pubppaa/US07_PUBCOMB.pep:*

2: /EMC_Celerra_SIDS3/prodata/2/pubppaa/US08_PUBCOMB.pep:*

3: /EMC_Celerra_SIDS3/prodata/2/pubppaa/US09_PUBCOMB.pep:*

4: /EMC_Celerra_SIDS3/prodata/2/pubppaa/US10_PUBCOMB.pep:*

5: /EMC_Celerra_SIDS3/prodata/2/pubppaa/US10_PUBCOMB.pep:*

6: /EMC_Celerra_SIDS3/prodata/2/pubppaa/US11_PUBCOMB.pep:*

SUMMARIES

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	DB ID	Description
1	2836	100.0	536	3	US-09-776-865-2
2	2836	100.0	536	5	US-10-823-506-8
3	2621	92.4	495	3	US-09-915-181A-7
4	2621	92.4	495	4	US-10-755-889-588
5	2621	92.4	495	5	US-10-823-506-2
6	2621	92.4	495	5	US-10-631-467-768
7	2329	82.1	495	3	US-09-776-865-4
8	2329	82.1	495	5	US-10-823-506-4
9	1698	59.9	495	5	US-10-823-506-10
10	1698	59.9	495	5	US-10-823-506-12
11	1521	53.6	284	4	US-10-264-049-2477
12	1451	51.2	272	4	US-10-264-237-2351
13	996.5	35.1	559	6	US-11-097-143-2895
14	992	35.0	502	6	US-11-097-143-8367
15	992	35.0	502	6	US-11-097-143-24411
16	987	34.8	493	4	US-10-369-493-5834
17	970	34.2	578	3	US-09-740-041-4
18	970	34.2	578	4	US-10-389-967-4
19	970	34.2	582	3	US-09-915-181A-4
20	970	34.2	582	3	US-10-205-331-7
21	970	34.2	582	5	US-10-734-731-12
22	970	34.2	582	5	US-10-734-731-14
23	970	34.2	582	5	US-10-807-500-12
24	970	34.2	582	5	US-10-807-500-14
25	965.5	34.0	582	5	US-10-734-731-10
26	965.5	34.0	582	5	US-10-807-500-10
27	955	33.7	589	3	US-09-740-041-2

28	955	33.7	589	4	US-10-389-967-2	Sequence 2, Appl1
29	955	33.7	589	5	US-10-499-731-28	Sequence 28, Appl1
30	945	33.3	601	5	US-10-499-731-46	Sequence 46, Appl1
31	941	33.2	560	5	US-10-734-731-2	Sequence 2, Appl1
32	941	33.2	560	5	US-10-807-500-2	Sequence 2, Appl1
33	941	33.2	560	5	US-10-756-149-5598	Sequence 5598, App
34	940.5	33.2	566	4	US-10-287-226-300	Sequence 300, App
35	940	33.1	560	4	US-10-314-790-5	GENERAL INFORMA
36	940	33.1	560	5	US-10-734-731-4	Sequence 4, Appl1
37	940	33.1	560	5	US-10-807-500-4	Sequence 4, Appl1
38	939	33.1	1294	4	US-10-286-115-1200	Sequence 1200, Ap
39	936	33.0	560	3	US-09-991-212A-4	Sequence 4, Appl1
40	936	33.0	560	3	US-09-915-181A-5	Sequence 5, Appl1
41	936	33.0	560	3	US-09-965-522-4	Sequence 4, Appl1
42	936	33.0	560	5	US-10-734-731-6	Sequence 6, Appl1
43	936	33.0	560	5	US-10-734-731-8	Sequence 8, Appl1
44	936	33.0	560	5	US-10-807-500-6	Sequence 6, Appl1
45	936	33.0	560	5	US-10-807-500-8	Sequence 8, Appl1

ALIGNMENTS

RESULT 1

US-09-776-865-2

Sequence 2, Application US/09776865

Patent No. US20020061846A1

GENERAL INFORMATION:

APPLICANT: Hellerqvist, Carl

TITLE OF INVENTION: Methods for Preventing or Attenuating Pathoangiogenic Conditions

FILE REFERENCE: 22100-0100 46126-252687

CURRENT APPLICATION NUMBER: US/09/776,865

CURRENT FILING DATE: 2001-02-02

PRIOR APPLICATION NUMBER: US 60/1179,870

PRIOR FILING DATE: 2000-02-02

NUMBER OF SEQ ID NOS: 4

SOFTWARE: PatentIn version 3.0

SEQ ID NO 2

LENGTH: 536

TYPE: PRT

ORGANISM: Homo sapiens

US-09-776-865-2

Query Match 100.0%; Score 2836; DB 3; Length 536;

Best Local Similarity 100.0%; Pred. No. 4.3e-249;

Matches 536; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAGMTTPRPVQPARPGFGLSGRRSLCQVASTPAHVGWMSPVRLDARNDBEESTDR 60

1 MAAGMTTPRPVQPARPGFGLSGRRSLCQVASTPAHVGWMSPVRLDARNDBEESTDR 60

QY 61 TPLLPAPRAEAPVCCSARYNLAIAFRGFYVYALRVNLVALVDMVDSNTTLEENRT 120

61 TPLLPAPRAEAPVCCSARYNLAIAFRGFYVYALRVNLVALVDMVDSNTTLEENRT 120

QY 61 TPLLPAPRAEAPVCCSARYNLAIAFRGFYVYALRVNLVALVDMVDSNTTLEENRT 120

61 TPLLPAPRAEAPVCCSARYNLAIAFRGFYVYALRVNLVALVDMVDSNTTLEENRT 120

QY 121 SKACPEHSAPIKYHNHNOTGKKYQMDAETQGMILGSEFFYGIITIQIPGGYASKIGKML 180

121 SKACPEHSAPIKYHNHNOTGKKYQMDAETQGMILGSEFFYGIITIQIPGGYASKIGKML 180

QY 121 SKACPEHSAPIKYHNHNOTGKKYQMDAETQGMILGSEFFYGIITIQIPGGYASKIGKML 180

121 SKACPEHSAPIKYHNHNOTGKKYQMDAETQGMILGSEFFYGIITIQIPGGYASKIGKML 180

QY 181 GFGILTAVALTLTPPIADIGVPLIVLRALBEGIGBGTVPAMHAMWSSVAPPLERSKL 240

181 GFGILTAVALTLTPPIADIGVPLIVLRALBEGIGBGTVPAMHAMWSSVAPPLERSKL 240

QY 181 GFGILTAVALTLTPPIADIGVPLIVLRALBEGIGBGTVPAMHAMWSSVAPPLERSKL 240

181 GFGILTAVALTLTPPIADIGVPLIVLRALBEGIGBGTVPAMHAMWSSVAPPLERSKL 240

QY 241 S1SYAQAQIGTVISLPSGIIICYYMMWTYVFYFGTIGTIFWLLMTLWVSDTQKHRIS 300

241 S1SYAQAQIGTVISLPSGIIICYYMMWTYVFYFGTIGTIFWLLMTLWVSDTQKHRIS 300

QY 241 S1SYAQAQIGTVISLPSGIIICYYMMWTYVFYFGTIGTIFWLLMTLWVSDTQKHRIS 300

241 S1SYAQAQIGTVISLPSGIIICYYMMWTYVFYFGTIGTIFWLLMTLWVSDTQKHRIS 300

QY 301 HYKEYITLSLRQSSOKSVPMVPLTKSLPLVAIVAHRSYVMTFTYTLTLPTWKKEI 360

301 HYKEYITLSLRQSSOKSVPMVPLTKSLPLVAIVAHRSYVMTFTYTLTLPTWKKEI 360

QY 301 HYKEYITLSLRQSSOKSVPMVPLTKSLPLVAIVAHRSYVMTFTYTLTLPTWKKEI 360

301 HYKEYITLSLRQSSOKSVPMVPLTKSLPLVAIVAHRSYVMTFTYTLTLPTWKKEI 360

QY 361 LRFNVQNGFLSLPYLGSWLCMILSGAADNLRAKNFSTLCVRRIFSLIGMIGPAVFL 420

361 LRFNVQNGFLSLPYLGSWLCMILSGAADNLRAKNFSTLCVRRIFSLIGMIGPAVFL 420

Db 361 LRFNVQENGFLSSLPLVGSWLCMLISGQAADNLRKKNFSTLCVRIRFSLIGMIGPAVFL 420
QY 421 VAAFGICDYSIAVAFLTITTTGGFCSSGFSINHLDIASVAGILLGINTPATIIPGMV 480
Db 421 VAAFGICDYSIAVAFLTITTTGGFCSSGFSINHLDIASVAGILLGINTPATIIPGMV 480
QY 481 GPVIAKSLTPDNTVGEQVTFYIAAINVGAIFFTLPFAKEVQNMALNDHGHHRH 536
Db 481 GPVIAKSLTPDNTVGEQVTFYIAAINVGAIFFTLPFAKEVQNMALNDHGHHRH 536

RESULT 2
US-10-823-506-8
Sequence 8, Application US/10823506
Publication No. US20050002931A1
GENERAL INFORMATION:
APPLICANT: Heileqvist, Carl
APPLICANT: Fu, Changlin
TITLE OF INVENTION: GBS Toxin Receptor
FILE REFERENCE: CARB-008/01US
CURRENT APPLICATION NUMBER: US/10/823,506
CURRENT FILING DATE: 2004-04-12
PRIOR APPLICATION NUMBER: US/09/359,167
PRIOR FILING DATE: 1999-07-21
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60-693,843
PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-22
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 8
LENGTH: 536
TYPE: PRF
ORGANISM: Homo sapiens
US-10-823-506-8

Query Match 100.0%; Score 2836; DB 5; Length 536;
Best Local Similarity 100.0%; Pred. No. 4,3e-249;
Matches 536; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAGANTPRPVQPARPGFGLSGRSLLCQVASTPAHVGMSPVVDLARNDEESTDR 60
Db 1 MAAGANTPRPVQPARPGFGLSGRSLLCQVASTPAHVGMSPVVDLARNDEESTDR 60
QY 61 TPLPGAPRAEAPVCCSARYNLAITLAFGFFIYVALRVNLVVALVMDVSDNTLEEDNR 120
Db 61 TPLPGAPRAEAPVCCSARYNLAITLAFGFFIYVALRVNLVVALVMDVSDNTLEEDNR 120
QY 121 SKACPEHSAPIKVHNQTKKKYQWDAETQWILGSPFYGYITTOIPGGYASAKTGGMML 180
Db 121 SKACPEHSAPIKVHNQTKKKYQWDAETQWILGSPFYGYITTOIPGGYASAKTGGMML 180
QY 181 GFGILGAVLTLFPIADIGVGPLIVLRALLEGIGEGVTPPAMHAMSSWAPLERSKL 240
Db 181 GFGILGAVLTLFPIADIGVGPLIVLRALLEGIGEGVTPPAMHAMSSWAPLERSKL 240
QY 241 SISVAGQLGTVISLPLSGIICYYMNTYVFFPGTIGIFWFLIMLVSDTPQKHRRIS 300
Db 241 SISVAGQLGTVISLPLSGIICYYMNTYVFFPGTIGIFWFLIMLVSDTPQKHRRIS 300
QY 301 HYEKEYLSSLRNQLSSQKSVWPVPIKSLPLMAIVVAHFSYNNTFYTLTLFTYKKEI 360
Db 301 HYEKEYLSSLRNQLSSQKSVWPVPIKSLPLMAIVVAHFSYNNTFYTLTLFTYKKEI 360
QY 361 LRFNVQENGFLSSLPLVGSWLCMLISGQAADNLRKKNFSTLCVRIRFSLIGMIGPAVFL 420
Db 361 LRFNVQENGFLSSLPLVGSWLCMLISGQAADNLRKKNFSTLCVRIRFSLIGMIGPAVFL 420
QY 421 VAAFGICDYSIAVAFLTITTTGGFCSSGFSINHLDIASVAGILLGINTPATIIPGMV 480
Db 421 VAAFGICDYSIAVAFLTITTTGGFCSSGFSINHLDIASVAGILLGINTPATIIPGMV 480
QY 481 GPVIAKSLTPDNTVGEQVTFYIAAINVGAIFFTLPFAKEVQNMALNDHGHHRH 536
Db 481 GPVIAKSLTPDNTVGEQVTFYIAAINVGAIFFTLPFAKEVQNMALNDHGHHRH 536

RESULT 3
US-09-915-181A-7
Sequence 7, Application US/09915181A
Patent No. US20020098473A1
GENERAL INFORMATION:
APPLICANT: EDWARDS, ROBERT
APPLICANT: BELLOCCHIO, ELIZABETH
APPLICANT: FREMEAUX, ROBERT
APPLICANT: REIMER, RICHARD
TITLE OF INVENTION: NOVEL GLUTAMATE TRANSPORTERS
FILE REFERENCE: 305T-932610US
CURRENT APPLICATION NUMBER: US/09/915,181A
CURRENT FILING DATE: 2002-03-26
PRIOR APPLICATION NUMBER: US 60/220,556
PRIOR FILING DATE: 2000-07-25
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PatentIn version 3.0
SEQ ID NO 7
LENGTH: 495
TYPE: PRF
ORGANISM: Homo sapiens
US-09-915-181A-7

Query Match 92.4%; Score 2621; DB 3; Length 495;
Best Local Similarity 100.0%; Pred. No. 1,5e-229;
Matches 495; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 MRSVPDLARNDEESTDRTPPLPGAPRAEAPVCCSARYNLAITLAFGFFIYVALRVNL 101
Db 1 MRSVPDLARNDEESTDRTPPLPGAPRAEAPVCCSARYNLAITLAFGFFIYVALRVNL 101
QY 102 SVLVVMDVSDNTLEEDNRSTKACPEHSAPIKVHNQTKKKYQWDAETQWILGSPFYGYI 161
Db 102 SVLVVMDVSDNTLEEDNRSTKACPEHSAPIKVHNQTKKKYQWDAETQWILGSPFYGYI 161
QY 162 ITQIPGGYASAKTGGMMLGFGILGAVLTLFPIADIGVGPLIVLRALLEGIGEGVTP 221
Db 162 ITQIPGGYASAKTGGMMLGFGILGAVLTLFPIADIGVGPLIVLRALLEGIGEGVTP 221
QY 222 AMHAMSSWAPLERSKLSTISVAGQLGTVISLPLSGIICYYMNTYVFFPGTIGIFW 281
Db 222 AMHAMSSWAPLERSKLSTISVAGQLGTVISLPLSGIICYYMNTYVFFPGTIGIFW 281
QY 282 FLIMLVSDTPQKHRRISHYEKEYLSSLRNQLSSQKSVWPVPIKSLPLMAIVVAHFS 341
Db 282 FLIMLVSDTPQKHRRISHYEKEYLSSLRNQLSSQKSVWPVPIKSLPLMAIVVAHFS 341
QY 342 YNNTFYTLTLFTYKKEILRFNVQENGFLSSLPLVGSWLCMLISGQAADNLRKKNFST 401
Db 342 YNNTFYTLTLFTYKKEILRFNVQENGFLSSLPLVGSWLCMLISGQAADNLRKKNFST 401
QY 402 LCVRRIFSLIGMIGPAVFLVAAFGICDYSIAVAFLTITTTGGFCSSGFSINHLDIAPS 461
Db 402 LCVRRIFSLIGMIGPAVFLVAAFGICDYSIAVAFLTITTTGGFCSSGFSINHLDIAPS 461
QY 462 YAGILLGINTPATIIPGMVGPVIAKSLTPDNTVGEQVTFYIAAINVGAIFFTLPFAK 521
Db 462 YAGILLGINTPATIIPGMVGPVIAKSLTPDNTVGEQVTFYIAAINVGAIFFTLPFAK 521
QY 522 EVQNMALNDHGHHRH 536
Db 522 EVQNMALNDHGHHRH 536
QY 481 GPVIAKSLTPDNTVGEQVTFYIAAINVGAIFFTLPFAK 480
Db 481 GPVIAKSLTPDNTVGEQVTFYIAAINVGAIFFTLPFAK 480

RESULT 4
US-10-755-889-588
Sequence 588, Application US/10755889
Publication No. US20040171823A1
GENERAL INFORMATION:
APPLICANT: Bristol-Myers Squibb Company
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-KB

```

; TITLE OF INVENTION: PATHWAY
; FILE REFERENCE: D0284 NP
; CURRENT APPLICATION NUMBER: US/10/755,889
; CURRENT FILING DATE: 2004-01-13
; PRIOR APPLICATION NUMBER: U.S. 60/440,068
; PRIOR FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: U.S. 60/469,757
; PRIOR FILING DATE: 2003-05-12
; NUMBER OF SEQ ID NOS: 823
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 588
; LENGTH: 495
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-755-889-588

```

```

Query Match      92.4%; Score 2621; DB 4; Length 495;
Best Local Similarity 100.0%; Pred. No. 1.5e-229;
Matches 495; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 42 MRSPPVRLARNDEESTDRTPLLPGAPRAEAPVCCSARYNLA ILAFGFEIYYALRVNL 101
DB 1 MRSPPVRLARNDEESTDRTPLLPGAPRAEAPVCCSARYNLA ILAFGFEIYYALRVNL 60
QY 102 SVALVDMVDSNTTLEDRNRTSKACPEHSAPIKVHHNQTGKKYQMDAETQGMILGSFFGYI 161
DB 61 SVALVDMVDSNTTLEDRNRTSKACPEHSAPIKVHHNQTGKKYQMDAETQGMILGSFFGYI 120
QY 162 ITQIPGGYVASKIGKMLGFGILGTAVLTLFTPIAADLGVPPLIVRALEGGEVTFP 221
DB 121 ITQIPGGYVASKIGKMLGFGILGTAVLTLFTPIAADLGVPPLIVRALEGGEVTFP 180
QY 222 AMHAMSSMAPPLERSKULSISYAGQLGTVISLPSGII CYNNMTYVYFFGTIGIFW 281
DB 181 AMHAMSSMAPPLERSKULSISYAGQLGTVISLPSGII CYNNMTYVYFFGTIGIFW 240
QY 282 FLIMWLVSDTPQKHRI SHYEKEYILSLRNQSSQKSVWPWPIKLSPLMLAI VVAHFS 341
DB 241 FLIMWLVSDTPQKHRI SHYEKEYILSLRNQSSQKSVWPWPIKLSPLMLAI VVAHFS 300
QY 342 YNNTFYTLTLPTMYKEILRFVVOENGFLSSLPYGSWLCMLISGQAADNLRKKNFST 401
DB 301 YNNTFYTLTLPTMYKEILRFVVOENGFLSSLPYGSWLCMLISGQAADNLRKKNFST 360
QY 402 LCVRRIFSLIGMIGPAVFLVAAFGICDYS LAVAFLITSTLGGFCSSGSGSINHLDIAPS 461
DB 361 LCVRRIFSLIGMIGPAVFLVAAFGICDYS LAVAFLITSTLGGFCSSGSGSINHLDIAPS 420
QY 462 YAGILGINTTPTATIPGMGVPVI AKSLTPNTVGENQTVFYIAAINVFGAIFFTLPAGK 521
DB 421 YAGILGINTTPTATIPGMGVPVI AKSLTPNTVGENQTVFYIAAINVFGAIFFTLPAGK 480
QY 522 EVONMALNDHGHGRH 536
DB 481 EVONMALNDHGHGRH 495

```

```

RESULT 5
US-10-823-506-2
; Sequence 2, Application US/10823506
; Publication No. US20050002931A1
; GENERAL INFORMATION:
; APPLICANT: Høllervæst, Carl
; APPLICANT: Fu, Changlin
; TITLE OF INVENTION: GBS Toxin Receptor
; FILE REFERENCE: CARB-008/01US
; CURRENT APPLICATION NUMBER: US/10/823,506
; CURRENT FILING DATE: 2004-04-12
; PRIOR APPLICATION NUMBER: US/09/359,167
; PRIOR FILING DATE: 1999-07-21
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60-693,843
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-22
; NUMBER OF SEQ ID NOS: 12

```

```

; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 495
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-823-506-2

```

```

Query Match      92.4%; Score 2621; DB 5; Length 495;
Best Local Similarity 100.0%; Pred. No. 1.5e-229;
Matches 495; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 42 MRSPPVRLARNDEESTDRTPLLPGAPRAEAPVCCSARYNLA ILAFGFEIYYALRVNL 101
DB 1 MRSPPVRLARNDEESTDRTPLLPGAPRAEAPVCCSARYNLA ILAFGFEIYYALRVNL 60
QY 102 SVALVDMVDSNTTLEDRNRTSKACPEHSAPIKVHHNQTGKKYQMDAETQGMILGSFFGYI 161
DB 61 SVALVDMVDSNTTLEDRNRTSKACPEHSAPIKVHHNQTGKKYQMDAETQGMILGSFFGYI 120
QY 162 ITQIPGGYVASKIGKMLGFGILGTAVLTLFTPIAADLGVPPLIVRALEGGEVTFP 221
DB 121 ITQIPGGYVASKIGKMLGFGILGTAVLTLFTPIAADLGVPPLIVRALEGGEVTFP 180
QY 222 AMHAMSSMAPPLERSKULSISYAGQLGTVISLPSGII CYNNMTYVYFFGTIGIFW 281
DB 181 AMHAMSSMAPPLERSKULSISYAGQLGTVISLPSGII CYNNMTYVYFFGTIGIFW 240
QY 282 FLIMWLVSDTPQKHRI SHYEKEYILSLRNQSSQKSVWPWPIKLSPLMLAI VVAHFS 341
DB 241 FLIMWLVSDTPQKHRI SHYEKEYILSLRNQSSQKSVWPWPIKLSPLMLAI VVAHFS 300
QY 342 YNNTFYTLTLPTMYKEILRFVVOENGFLSSLPYGSWLCMLISGQAADNLRKKNFST 401
DB 301 YNNTFYTLTLPTMYKEILRFVVOENGFLSSLPYGSWLCMLISGQAADNLRKKNFST 360
QY 402 LCVRRIFSLIGMIGPAVFLVAAFGICDYS LAVAFLITSTLGGFCSSGSGSINHLDIAPS 461
DB 361 LCVRRIFSLIGMIGPAVFLVAAFGICDYS LAVAFLITSTLGGFCSSGSGSINHLDIAPS 420
QY 462 YAGILGINTTPTATIPGMGVPVI AKSLTPNTVGENQTVFYIAAINVFGAIFFTLPAGK 521
DB 421 YAGILGINTTPTATIPGMGVPVI AKSLTPNTVGENQTVFYIAAINVFGAIFFTLPAGK 480
QY 522 EVONMALNDHGHGRH 536
DB 481 EVONMALNDHGHGRH 495

```

```

RESULT 6
US-10-631-467-768
; Sequence 768, Application US/10631467
; Publication No. US20050208496A1
; GENERAL INFORMATION:
; APPLICANT: Genox Research Inc.
; TITLE OF INVENTION: Method for testing for bronchial asthma, or chronic obstructive
; FILE REFERENCE: 3462.1005-000
; CURRENT APPLICATION NUMBER: US/10/631,467
; CURRENT FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: JP 2003-077212
; PRIOR FILING DATE: 2003-03-20
; PRIOR APPLICATION NUMBER: JP 2002-229312
; PRIOR FILING DATE: 2002-08-06
; NUMBER OF SEQ ID NOS: 2086
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 768
; LENGTH: 495
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-631-467-768

```

```

Query Match      92.4%; Score 2621; DB 5; Length 495;
Best Local Similarity 100.0%; Pred. No. 1.5e-229;

```

Matches	Conservative	0	Mismatches	0	Indels	0	Gaps	0
Qy	42	MRSPIRDLARNDEESTDRTPLPGAPRAEAPVCCSARYNLALIAFFGFTIVYALRVNL	101					
Db	1	MRSPIRDLARNDEESTDRTPLPGAPRAEAPVCCSARYNLALIAFFGFTIVYALRVNL	60					
Qy	102	SVALVDWDSNTTLEDNRTSKACPEHSAPIKVHNNOTGKKYQMDAETQGMILGSPFYGYI	161					
Db	61	SVALVDWDSNTTLEDNRTSKACPEHSAPIKVHNNOTGKKYQMDAETQGMILGSPFYGYI	120					
Qy	162	ITQIPGVYASIKGGKMLGFGILGTAVLTFTPIAADVGVPILYLRALBEGSGVTFP	221					
Db	121	ITQIPGVYASIKGGKMLGFGILGTAVLTFTPIAADVGVPILYLRALBEGSGVTFP	180					
Qy	222	AMHAMSSWAPPLERSKULSLISYAGAQGTIVISLPSGIIICYMMNTVYVFYFGTIGLPW	281					
Db	181	AMHAMSSWAPPLERSKULSLISYAGAQGTIVISLPSGIIICYMMNTVYVFYFGTIGLPW	240					
Qy	282	FLMWIWLVSDFPQKHRIISHYEKEYILSSLRNQLSQKSVWPVPIKSLPLMAIYVAHS	341					
Db	241	FLMWIWLVSDFPQKHRIISHYEKEYILSSLRNQLSQKSVWPVPIKSLPLMAIYVAHS	300					
Qy	342	YNWTEYTTLLTLPTMKETLRFNVOENGSLSPYIGSLCILSGQADNLRKXNFT	401					
Db	301	YNWTEYTTLLTLPTMKETLRFNVOENGSLSPYIGSLCILSGQADNLRKXNFT	360					
Qy	402	LCVRRIFSLIGMIGVAPFLVAAGFTGCDYSLAVAFLLTSTLGGFCSSGFSINHLDIAPS	461					
Db	361	LCVRRIFSLIGMIGVAPFLVAAGFTGCDYSLAVAFLLTSTLGGFCSSGFSINHLDIAPS	420					
Qy	462	YAGIILGTTNFTATIPGMVGVIAKSLTPDNVTGEMQVFIYIAAINVGAIFFTLPFAKG	521					
Db	421	YAGIILGTTNFTATIPGMVGVIAKSLTPDNVTGEMQVFIYIAAINVGAIFFTLPFAKG	480					
Qy	522	EVQNMALNDHGHRRH	536					
Db	481	EVQNMALNDHGHRRH	495					
RESULT 7								
US-09-776-865-4								
Sequence 4, Application US/09776865								
Patent No. US20020061846A1								
GENERAL INFORMATION:								
APPLICANT: Hellerqvist, Carl								
TITLE OF INVENTION: Methods for Preventing or Attenuating Pathoangiogenic Conditions								
FILE REFERENCE: 22100-0100 46126-252687								
CURRENT APPLICATION NUMBER: US/09/776,865								
CURRENT FILING DATE: 2001-02-02								
PRIOR APPLICATION NUMBER: US 60/179,870								
PRIOR FILING DATE: 2000-02-02								
NUMBER OF SEQ ID NOS: 4								
SOFTWARE: PatentIn version 3.0								
SEQ ID NO 4								
LENGTH: 495								
TYPE: PRT								
ORGANISM: Ovis sp.								
US-09-776-865-4								
Query Match								
Best Local Similarity 82.1%; Score 2329; DB 3; Length 495;								
Matches 427; Conservative 39; Mismatches 29; Indels 0; Gaps 0;								
Qy	42	MRSPIRDLARNDEESTDRTPLPGAPRAEAPVCCSARYNLALIAFFGFTIVYALRVNL	101					
Db	1	MRSPIRDLARNDEESTDRTPLPGAPRAEAPVCCSARYNLALIAFFGFTIVYALRVNL	60					
Qy	102	SVALVDWDSNTTLEDNRTSKACPEHSAPIKVHNNOTGKKYQMDAETQGMILGSPFYGYI	161					
Db	61	SVALVDWDSNTTLEDNRTSKACPEHSAPIKVHNNOTGKKYQMDAETQGMILGSPFYGYI	120					
Qy	162	ITQIPGVYASIKGGKMLGFGILGTAVLTFTPIAADVGVPIL						

Db	121	ITDIPGGYVASSGGKLLIGGIPATATLFTFLPLADFGVGLVALRALEGGGEVTP	180
QY	222	AMAMAMSSAAPLESKLLISISYAGAOGTVISLPISGIIICYNMNTVYVFFGTGIFW	281
Db	181	AMAMAMSSAAPLESKLLISISYAGAOGTVISLPISGVICYNMNTVYVFFGIVGIW	240
QY	282	FLMTLVSDTPOQKHRIASHPEKEIYLLSSLNQSSSKSVPMWPIIKSLPLMAIVAAHS	341
Db	241	FLMTICLVSDTPEHKTITPYEKEIYLLSSLNQSSSKSVPMWPIIMLSLPLMAIVAAHS	300
QY	342	YMTPEYTLTLTLPTWMEKILRENVQENGLSSLPYLGSMVCMILSGAADNLAKNNFST	401
Db	301	YMTPEYTLTLTLPTWMEKILRENVQENGLSSLPYLGSMVCMILSGAADNLAKNNFST	360
QY	402	LCVRRIFSLIGMIGPAVLVAGFGICDYSLAVAFLLTSTLLGGFCSSGFSINHLDIAPS	461
Db	361	LWRRVRFSLIGMIGPAIFLVAGFGICDYSLAVAFLLTSTLLGGFCSSGFSINHLDIAPS	420
QY	462	YAGILLGINTPATITPGWVGPIIASLTPDNTVEMQTVFIAMAINVGAIFFTLPARG	521
Db	421	YAGILLGINTPATITPGMIGPIIASLTPENTIGEMQTVFCIAAINVGAIFFTLPARG	480
QY	522	EYONMALNDHGHHRH 536	
Db	481	EYONMALSDHQHRN 495	
RESULT 8			
US-10-823-506-4			
; Sequence 4, Application US/10823506			
; Publication No. US20050002931A1			
; GENERAL INFORMATION:			
; APPLICANT: Hellegqvist, Carl			
; APPLICANT: Fu, Changlin			
; TITLE OF INVENTION: GBS Toxin Receptor			
; FILE REFERENCE: CARB-008/01US			
; CURRENT APPLICATION NUMBER: US/10/823,506			
; CURRENT FILING DATE: 2004-04-12			
; PRIOR APPLICATION NUMBER: US/09/359,167			
; PRIOR FILING DATE: 1999-07-21			
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60-693,843			
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-22			
; NUMBER OF SEQ ID NOS: 12			
; SOFTWARE: PatentIn Ver. 2.0			
; SEQ ID NO 4			
; LENGTH: 495			
; TYPE: PRT			
; ORGANISM: Ovis sp.			
US-10-823-506-4			
Query Match 82.1%; Score 2329; DB 5; Length 495;			
Best Local Similarity 86.3%; Pred. No. 5,9e-203;			
Matches 427; Conservative 39; Mismatches 29; Indels 0; Gaps 0			
QY	42	MRSPPVDLARNDGEESTDRTPILPAPAPAEAPVCCSRRYVLLALAFEGGFYIYALRVNL	101
Db	1	MKSPVSDLLAPSDGEESSDRTPILQAPAPAEAPVCCSRRYVLAFLSFGGFVLYLSRVNL	60
QY	102	SVLVLMVDSNTTLEENRTSKACPEHSAPIRKVNHNQTKKQMPAEFOGAILDSFFGYI	161
Db	61	SVLVLMVDSNTTAKDNRTSKYCAHSAPIRKVLHNQTKKRYMDAEFOGAILDSFFGYI	120
QY	162	ITDIPGGYVASSKIGKMLIGFGLTAVLLFTPLADLGVPLIVLRALEGGGEVTP	221
Db	121	ITDIPGGYVASSGGKLLIGGIPATATLFTFLPLADFGVGLVALRALEGGGEVTP	180
QY	222	AMAMAMSSAAPLESKLLISISYAGAOGTVISLPISGIIICYNMNTVYVFFGTGIFW	281
Db	181	AMAMAMSSAAPLESKLLISISYAGAOGTVISLPISGVICYNMNTVYVFFGIVGIW	240
QY	282	FLMTLVSDTPOQKHRIASHPEKEIYLLSSLNQSSSKSVPMWPIIKSLPLMAIVAAHS	341
Db	241	FLMTICLVSDTPEHKTITPYEKEIYLLSSLNQSSSKSVPMWPIIMLSLPLMAIVAAHS	300
QY	342	YMTPEYTLTLTLPTWMEKILRENVQENGLSSLPYLGSMVCMILSGAADNLAKNNFST	401
Db	301	YMTPEYTLTLTLPTWMEKILRENVQENGLSSLPYLGSMVCMILSGAADNLAKNNFST	360
QY	402	LCVRRIFSLIGMIGPAVLVAGFGICDYSLAVAFLLTSTLLGGFCSSGFSINHLDIAPS	461
Db	361	LWRRVRFSLIGMIGPAIFLVAGFGICDYSLAVAFLLTSTLLGGFCSSGFSINHLDIAPS	420
QY	462	YAGILLGINTPATITPGWVGPIIASLTPDNTVEMQTVFIAMAINVGAIFFTLPARG	521
Db	421	YAGILLGINTPATITPGMIGPIIASLTPENTIGEMQTVFCIAAINVGAIFFTLPARG	480
QY	522	EYONMALNDHGHHRH 536	
Db	481	EYONMALSDHQHRN 495	

Qy	342	NNWTFYTLTLTLPTMYKELTRFNVOENGSLSSPYGSMCMLSQOADNLEAKNNFST	401
Db	301	YNWTFYTLTLTLPTMYKEVLRNENIQENGFLSNAPYICMCMLSQOADNLEAKNNFST	360
Qy	402	LCVRRI FSLIGMIGPAVLVAAGFICDYS LAVAFLTISTTGGFSSGFSINHLDIAPs	461
Db	361	LMVRVVFSLIGMIGPAFLVAAGFICDYS LAVAFLTISTTGGFSSGFSINHLDIAPs	420
Qy	462	YAGILLGITNTFATITPGMGPVIAKSLTETDNTYGEWQTYFYIAAINVFGALFTTLPAKG	521
Db	421	YAGILLGITNTFATITPGMGPVIARBLTPENTNIGEWQTFECIAAAINVFGALFTTLPAKG	480
Qy	522	EYQNMALNDHGHGRH	536
Db	481	EYQNMALISDHQSHRN	495

```

RESULT 9
US-10-823-506-10
; Sequence 10, Application US/10823506
; Publication No. US2005000231A1
; GENERAL INFORMATION:
; APPLICANT: Hellerqvist, Carl
; APPLICANT: Fu, Changlin
; TITLE OF INVENTION: G8S Toxin Receptor
; FILE REFERENCE: CARB-008/01us
; CURRENT APPLICATION NUMBER: US/10/823,506
; CURRENT FILING DATE: 2004-04-12
; PRIOR APPLICATION NUMBER: US/09/359,167
; PRIOR FILING DATE: 1999-07-21
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60-693,843
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-22
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 495
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION:
; US-10-823-506-10

```

```

Query Match      59.9%; Score 1698; DB 5; Length 495;
Query Similarity 67.8%; Pred. No. 1.7e-145;
Matches 335; Conservative 0; Mismatches 159; Indels 0; Gaps 0

QY      42  MRSPIVDIARNDGEESTDRTPLPAPAPAAEAPVCCSARYNALAIAPFGFFIVALARNL 101
Db       1  MXXPXDXDXXXGGEFXDXDXXXXXXAXXEXEAPXCSCSARYNAXALPFGFFXXYLAXNL 60

QY      102 SVALVDWVDSNTLTEDNRTSKACPEHSAPIKVHNQTKKYQWDAETGMIIGSFYGYI 161
Db       61  XVXXVXXDXDSTXXDXDNRXSXXCXEHSAPIKVXXQTKKYWDDETGMILXFXGYI 120

QY      162 ITQIPGVYASKIGSKMLIGFQILGTAVULTFTPIAADLGVGPLVDALELGSGVTP 221
Db       121  ITXIPGVYASXGXXKLGXIGIXXXAXXTFTPPAADXGXGXXXXLXALELGGXGTXP 180

QY      222 AMHAMWSMAPLEERSKLLSTISYAGNAGTVISLPSGICVYMMWTVYFYPFGTIGTFW 281
Db       181  AMHAMWSMAPLEERSKLLXIXIYAPAKIGYASLPSGICVYMMWTVYFXXFXKXGXW 240

QY      282 FLIWIWLSDFPOKHRIISHEKEYILSLRNQSSQKSVPVPIKLSLPLMAIVAHFS 341
Db       241  FXMXIXLVSKFPXHKHXXXXXXEKKXILSLRNQSSQKSVPVXXXXKLLPLMAIXVAFS 300

QY      342 YNMTFTYTLTLPLTYMKEILRFNVQENGLSLPYLGMLCMIISGQADNIRAXWNST 401
Db       301  YNMTFTYTLTLPLTYMKXXLRFNVQENGLSXPLXXMLCMIISGQADNIRAXWNST 360

QY      402 LCVRIRISFLIGMIGAVFLVAAAGFIGCDYSLAVALTITSTLGGCCSSGGSINLHDIPS 461
Db       402  LCVRIRISFLIGMIGAVFLVAAAGFIGCDYSLAVALTITSTLGGCCSSGGSINLHDIPS 461

```

Db 361 XXVXRBSLIXMIGPXXFLXXXXXXXXGCOPYXLXVFLXI STLGGFCSSGFSINHLXIAPS 420

Qy 462 YACILGINTPATITGCMGPIVIASTLPDNTGSEMOVFYIAAIIINFGAIFETLPAGK 521

Db 421 YAGXLGIIINXFTITIGMGPIXIXSXPTXNTGEMOXFKXAAAXXNFGAIFETLPAGK 480

Qy 522 EVQNMALNDHHGHR 535

Db 481 EKQNMXXXDXHGHR 494

```

RESULT 10
US-10-823-506-12
: Sequence 12, Application US/10823506
: Publication No. US20050002931A1
: GENERAL INFORMATION:
: APPLICANT: Hellerqvist, Carl
: APPLICANT: Fu, Changlin
: TITLE OF INVENTION: GBS Toxin Receptor
: FILE REFERENCE: CARB-008/01US
: CURRENT APPLICATION NUMBER: US/10/823,506
: CURRENT FILING DATE: 2004-04-12
: PRIOR APPLICATION NUMBER: US/09/359,167
: PRIOR FILING DATE: 1999-07-21
: PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60-693,843
: PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-22
: NUMBER OF SEQ ID NOS: 12
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 12
: LENGTH: 495
: TYPE: PRT
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: :
US-10-823-506-12

```

Query Match	59.9%	Score 1698	DB 5	Length 495
Best Local Similarity	67.8%	Pred. No. 1.7e-145		
Matches 335	Conservative 0	Mismatches 159	Indels 0	Gaps 0
QY	42	MRSFVRLARNDGEESTDRTPFLPGAPRAEAPVCCSARVNLALAFPGFPIVVALRVNL	101	
Db	1	MXHXVHXAXXGXGEBXXDRXXXXXXXXXXARXEAAPCCSARVNXAXLAFPGFPIYXLYXNL	60	
QY	102	SVALVDWDSNTLTLEDNRTSFACPEHSAPIKVHNHQTGKKYQMDAETGAILGSEFYIYI	161	
Db	61	XVXXVXXMXDXSTXTXXDNRXSXXCXEHSAPIKVXXXTGKXKXYMDAETGAILXXFXFYGI	120	
QY	162	ITQIPGVYVASKTGGKMLFPGILGTAVLTFTFIADLTGVPILVRALEGLEGVTFP	221	
Db	121	ITXIPGVYVASXXGKXXXXLGXIXXXAXTTLFTFYAADXGKXXXXXIXALEXIDEGKXP	180	
QY	222	AMHAMSSMAPLEBSKILSISVGAQDGTIVSLPLSGIICYYNMNTYVFYFPGHIGFW	281	
Db	181	AMHAMSSMAPLEBSKILXIXIYAGALIGTVSLPLSGXICYNMNTYVFYFXXFXGXW	240	
QY	282	FLMIWLVSDTPQGHKRI SHYEKEYIISLRNQLSQQSVWPVBIKSLPLMAIVAHFS	341	
Db	241	FXMIXIVXSTPPXHXKXXXXXXEKKXIIISLXNQSSQKSVWXXXXXKCLPLMAXVXAFS	300	
QY	342	YNNFTFTYLLTLPLTYKELLRFNVQENGFSLSPLYLCSWLCMILSGAADNLRKKNFST	401	
Db	301	YNNFTFYTXKLLPLTYMKXXLRFNVQENGFLSXKXYLXKXWCMILXGQAADNLRAXNMFST	360	
QY	402	LCVARIISLGMIGPAVFLVAAGRIIGDYSIAVAFLTISTTLGGFCSSGFSINHLDIAPS	461	
Db	361	XXVVRKFSLIXMIGPXXFLXXXXXXCDYKIXVAFLTISTLGGFCSSGFSINHLXIAPS	420	
QY	462	YAGILLETITTFATIPGMVGPVIAKSLTLPNTTVGEMOTVFYIAAIVFCAIFPTLPAKG	521	
Db	421	YAGLLLETITXKFAITIXGMKGPXIXXKSTPNTXGEMGXFXKXAAANXFGAIFPTLPAKG	480	
QY	522	EVQNMALNDHHGHR	535	

Db 481 EXONXXXXXGHR 494

RESULT 11

US-10-264-049-2477
; Sequence 2477, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
; APPLICANT: Birste et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA133P1
; CURRENT APPLICATION NUMBER: US/10/264,049
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 2477
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-264-049-2477

Query Match 53.6%; Score 1521; DB 4; Length 284;
Best Local Similarity 100.0%; Pred. No. 1,1e-129;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 253 ISLPLSGIICYYMMNTYVFYFEGTIGIFWFLMIMLVSDTPQKHRI SHYEKEYILSLR 312
Db 1 ISLPLSGIICYYMMNTYVFYFEGTIGIFWFLMIMLVSDTPQKHRI SHYEKEYILSLR 60
QY 313 NQSSQKSVMPVPIIKSLPLMAIIVAHFSYNTFTYLLTLPYMKELIRNVQENGFLS 372
Db 61 NQSSQKSVMPVPIIKSLPLMAIIVAHFSYNTFTYLLTLPYMKELIRNVQENGFLS 120
QY 373 SLPLSGWLCMILSGQADNLRANKNFSTLCVRRI FSLIGMIGPAVFLVAGFTGCDYSL 432
Db 121 SLPLSGWLCMILSGQADNLRANKNFSTLCVRRI FSLIGMIGPAVFLVAGFTGCDYSL 180
QY 433 AVAFLTSTLTGFCSSGFSINHLDIAPSYAGILGINTPATIPGMVGPVIAKSLTPDN 492
Db 181 AVAFLTSTLTGFCSSGFSINHLDIAPSYAGILGINTPATIPGMVGPVIAKSLTPDN 240
QY 493 TVGEMQTVFYIAAIVNFGAIFFTLFAKGEVQNMALNDHGHHRH 536
Db 241 TVGEMQTVFYIAAIVNFGAIFFTLFAKGEVQNMALNDHGHHRH 284

RESULT 12

US-10-264-237-2351
; Sequence 2351, Application US/10264237
; Publication No. US20040009491A1
; GENERAL INFORMATION:
; APPLICANT: Birste et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA131P1
; CURRENT APPLICATION NUMBER: US/10/264,237
; PRIOR APPLICATION NUMBER: PCT/US01/16450
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205,515
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 2876
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 2351
; LENGTH: 272
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC_FEATURE

; LOCATION: (55)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
US-10-264-237-2351

Query Match 51.2%; Score 1451; DB 4; Length 272;
Best Local Similarity 99.6%; Pred. No. 2,4e-123;
Matches 271; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 265 MNMTYVFYFEGTIGIFWFLMIMLVSDTPQKHRI SHYEKEYILSLRNQSSQKSVMPW 324
Db 1 MNMTYVFYFEGTIGIFWFLMIMLVSDTPQKHRI SHYEKEYILSLRNQSSQKSVMPW 60
QY 325 PILKSLPLMAIIVAHFSYNTFTYLLTLPYMKELIRNVQENGFLSLPYLGSWLCM1 384
Db 61 PILKSLPLMAIIVAHFSYNTFTYLLTLPYMKELIRNVQENGFLSLPYLGSWLCM1 120
QY 385 LSGQADNLRANKNFSTLCVRRI FSLIGMIGPAVFLVAGFTGCDYSLAVAFLLSTTLG 444
Db 121 LSGQADNLRANKNFSTLCVRRI FSLIGMIGPAVFLVAGFTGCDYSLAVAFLLSTTLG 180
QY 445 GFCSSGFSINHLDIAPSYAGILGINTPATIPGMVGPVIAKSLTPDN TVGEMQTVFYIA 504
Db 181 GFCSSGFSINHLDIAPSYAGILGINTPATIPGMVGPVIAKSLTPDN TVGEMQTVFYIA 240
QY 505 AAINVFGAIFFTLFAKGEVQNMALNDHGHHRH 536
Db 241 AAINVFGAIFFTLFAKGEVQNMALNDHGHHRH 272

RESULT 13

US-11-097-143-2895
; Sequence 2895, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: PaedSeq for Windows Version 4.0
; SEQ ID NO 2895
; LENGTH: 559
; TYPE: PRT
; ORGANISM: DROSOPHILA

US-11-097-143-2895

Query Match 35.1%; Score 996.5; DB 6; Length 559;
Best Local Similarity 39.8%; Pred. No. 1,6e-81;
Matches 208; Conservative 91; Mismatches 171; Indels 53; Gaps 10;

QY 51 RNDEESTDR-----TPLLPGAPRAEAPVVC-CSARYNLAIALAFGPFIVYALRYVLS 102
Db 31 RSDDEADDERAFCGGERPLIRSSGAABENHGCSPKTRHIFGFWGFLGFAVVYIMRVNLS 90

QY 103 VALDMDV-----SNTLEDNRKTSKACPEHSAPIKYHNN-----QTKKKYQMDAETQG 150
 Db 91 VAIAMVWQTAIPHNSNSSVIDTDT---CP-LPAF---HHNSGDPHQKGEVWMBEATQG 143
 QY 151 WILSFFPGYIITQIPGVYASKIGKMLGFGILGTAVLTFPIAADLGVPLIVRA 210
 Db 144 LVLSFFGYVLTQVPGGRMAELYGKKIYGVGLITAVFTLITPLAAMDLPLILVRI 203
 QY 211 LEGGEGTTPPAMAMSSMAPLERSKLISYVAGLQGVYISPLSGIIC---YNNW 267
 Db 204 LEGMBEGVTPPAMHMLHMLPPLERKRPALIVAGSNGITVISMPLGMLCSLDFLGW 263
 QY 268 TYVFFPGTIGIFWELMWIWLVSPTPOKHKRISHYEKEYILSSLRNQ----- 314
 Db 264 PSAFYIFGLGLIMWIAMMYLVYDKSDHPRISESERIERSLQVQLINDLAABEE 323
 QY 315 -----LSQKSVPVPLIKSLPLMAIVAHFSYMTFTYTLTLPTWKELLRPNV 365
 Db 324 EGODEVSLRAPPREEPIPMSSLLTSVPLMAILLTQCGQMAFYQTLETPTYSNIIHPDI 383
 QY 366 QENGFLSLPYLGSGLCMILSGOANDLRKKNFSTLCVRRIPLSLGIMGPAVPLVAGF 425
 Db 384 QSNMLNANVPYLTSEFVGIACSALDMLARYISLNSYKLMNTVASVPSLGLITIIY 443
 QY 426 IGCYSLAVAFLLTSTLGFCSGFSINHLDIAPSVAGIILGINTPATTPGMWGP-VI 484
 Db 444 VCGMWVMTFPLAGVSGFAGVAGNQNQHIALSRVAGTMYGITSANICGFLAPYI 503
 QY 485 AKSLTPDNTGEMQTVFYIAAIVFGAIFFTLPAKGEVONWA 527
 Db 504 GLIINHRETLQWHLVFWLAAGLNIAGNFYILIFASAEQGSWS 546

RESULT 14
 US-11-097-143-8367
 ; Sequence 8367, Application US/11097143
 ; Publication No. US20050208558A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Venter, J. Craig
 ; APPLICANT: et al.
 ; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
 ; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
 ; FILE REFERENCE: CLO00728
 ; CURRENT APPLICATION NUMBER: US/11/097,143
 ; CURRENT FILING DATE: 2005-04-04
 ; PRIOR APPLICATION NUMBER: 60/157,832
 ; PRIOR FILING DATE: 1999-10-05
 ; PRIOR APPLICATION NUMBER: 60/160,191
 ; PRIOR FILING DATE: 1999-10-19
 ; PRIOR APPLICATION NUMBER: 60/161,932
 ; PRIOR FILING DATE: 1999-10-28
 ; PRIOR APPLICATION NUMBER: 60/164,769
 ; PRIOR FILING DATE: 1999-11-12
 ; PRIOR APPLICATION NUMBER: 60/173,383
 ; PRIOR FILING DATE: 1999-12-28
 ; PRIOR APPLICATION NUMBER: 60/175,693
 ; PRIOR FILING DATE: 2000-01-12
 ; PRIOR APPLICATION NUMBER: 60/184,831
 ; PRIOR FILING DATE: 2000-02-24
 ; PRIOR APPLICATION NUMBER: 60/191,637
 ; PRIOR FILING DATE: 2000-03-23
 ; NUMBER OF SEQ ID NOS: 43008
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 8367
 ; LENGTH: 502
 ; TYPE: PRT
 ; ORGANISM: DROSOPHILA
 US-11-097-143-8367
 Query Match 35.0%; Score 992; DB 6; Length 502;
 Best Local Similarity 42.4%; Pred. No. 3.5e-81;

Matches 190; Conservative 93; Mismatches 149; Indels 16; Gaps 3;
 QY 80 RYNLIAPFGFPIYALRVNLASVALVMDVNSNTLEDNRKTSKACPEHSAPIKYHNNQTG 139
 Db 38 RYIVVILAFPGFENVYSIRVNLVAIVAMTENVYFD-----AGNNYSQO-- 83
 QY 140 KKYOMDAETQGMILSGFFGYIITQIPGVYASKIGKMLGFGILGTAVLTFPIAAD 199
 Db 84 -DFPMSKQKGLILSSFFGYIITQIPGVYASKIGKMLGFGILGTAVLTFPIAAD 142
 QY 200 LGVGPLIVRLLEGGEVTPPAMAMSSMAPLERSKLISYVAGLQGVYISPLSG 259
 Db 143 HSLWFLFVRRIIEGFFBEVTPPGIHAVARMSPLERSRMASIFAGVAGTIVAMPSCG 202
 QY 260 IICVMMNTVYFPGTIGIFWELMWIWLVSPTPOKHKRISHYEKEYILSSLRNQLSQX 319
 Db 203 FLATKGMESVYFVGITIGVITWLVFVAGSPLDRFCSKEEDYIQTITGVYGSXIV 262
 QY 320 SVPVPLIKSLPLMAIVAHFSYMTFTYTLTLPTWKELLRPNVQENGFLSLPYLGS 379
 Db 263 KHPRAIFTSMPFALMAHSHSEMGFTTLTQLPSFLRDTLNFDLGKGLSAVPLAM 322
 QY 380 WLCMILSGOANDLRKKNFSTLCVRRIPLSLGIMGPAVPLVAGFICGYSLAVAFIT 439
 Db 323 GILAVSGYLDWLDQVKGIMTITQVRRNFNGCAFIAQTFWMLTAYL-LDPTWSVSLTI 381
 QY 440 STTLGFCSSGFSINHLDIAPSVAGIILGINTPATTPGMWGPVIAKSLTPDNTGEMQ 499
 Db 382 AVIGAFPMWGSFANVHLDIAPQHASVLMGINTPATTPGIIVSLTLGVTYVNTQSDWKRI 441
 QY 500 VFYIAAIVFGAIFFTLPAKGEVONWA 527
 Db 442 IFIISAGIYLVGCVIYWFYCSGLQEWNA 469

RESULT 15
 US-11-097-143-24411
 ; Sequence 24411, Application US/11097143
 ; Publication No. US20050208558A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Venter, J. Craig
 ; APPLICANT: et al.
 ; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
 ; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
 ; FILE REFERENCE: CLO00728
 ; CURRENT APPLICATION NUMBER: US/11/097,143
 ; CURRENT FILING DATE: 2005-04-04
 ; PRIOR APPLICATION NUMBER: 60/157,832
 ; PRIOR FILING DATE: 1999-10-05
 ; PRIOR APPLICATION NUMBER: 60/160,191
 ; PRIOR FILING DATE: 1999-10-19
 ; PRIOR APPLICATION NUMBER: 60/161,932
 ; PRIOR FILING DATE: 1999-10-28
 ; PRIOR APPLICATION NUMBER: 60/164,769
 ; PRIOR FILING DATE: 1999-11-12
 ; PRIOR APPLICATION NUMBER: 60/173,383
 ; PRIOR FILING DATE: 1999-12-28
 ; PRIOR APPLICATION NUMBER: 60/175,693
 ; PRIOR FILING DATE: 2000-01-12
 ; PRIOR APPLICATION NUMBER: 60/184,831
 ; PRIOR FILING DATE: 2000-02-24
 ; PRIOR APPLICATION NUMBER: 60/191,637
 ; PRIOR FILING DATE: 2000-03-23
 ; NUMBER OF SEQ ID NOS: 43008
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 24411
 ; LENGTH: 502
 ; TYPE: PRT
 ; ORGANISM: DROSOPHILA
 US-11-097-143-24411
 Query Match 35.0%; Score 992; DB 6; Length 502;

Best Local Similarity 42.4%; Pred. No. 3.5e-81;
Matches 190; Conservative 93; Mismatches 149; Indels 16; Gaps 3;

```

QY      80 RYNLAIIAEFGFPIVALRVNLVSALVDVDSNTTLEDNRTSKACPEHSAPIKVHHNQIG 139
      ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      38 RYIVVLLAFPGFNNVSLRVNLVAIVAMTENRTVFD-----ADGNVSYQQ-- 83
QY      140 KKYQMDAEIQTGMLGSEFFGYITITQIPGIVASKIGKMLGFGILGTAVLTLETPRIAD 139
      :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db      84 -DEPWDSKQKGLISSEFFGYITITQIPGIVASKIGKMLGFGILGTAVLTLETPRIAD 142
QY      200 LGVGPILIVRALLEGVGTPEPRAMHAWSSMAPLEERSKULSISYAGAQIGTVISLPLSG 259
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      143 HSLMFLEFVRIIEGFEQVTFPGIHAWARMSPLEERSKVASIAFAGNAGTVVAMPCSG 202
QY      260 IICYYMMWTVVFFYFPGTIGIFWFLMIWLVSDTPQKHRI SHYEKEYILSLRNQLSSQK 319
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      203 FLATKXGMESEVFYFPGTIGIWIWLVFVKAGBELDRFCSKECDYIQKTI GYVGSKHV 262
QY      320 SVPWVPLKSLPLMAIVVAHFSYNWTFYITLTLPTMKELRNVOENGFSLSPYIGS 379
      ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      263 KHPWRAIFTSMPFYAIWASHFSEWGFYITLTLPSFLRDTLNFDLGKTGILSAVPYLA 322
QY      380 WLCMIISGQADNLRKAMNFTLCVRRIFSLIGWIGPAVFLVAAFGICDYSIAVAEFTI 439
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      323 GILLAVSGYLADMLQVKGITTTQVRRNFCAGFLAQTFEMMLTAYL-LDPTWSVSLTI 381
QY      440 STTLGGFCSSGFSINHLDIAPSYAGILLGITNTFATIPGMVGPVIAKSLTPDNTVGEWQT 499
      : ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      382 AVGLGAFWMSGFPAVNHLDIAQHASVLMGIGNTFATIPGI VSPLLTGTGVVNTQTSDEMRI 441
QY      500 VFYIAAIAINFGAIEFTLFAKGEVQNNWA 527
      :||: :||: :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db      442 IFFISAGIYLVGCIVWFYCSGDLQEWNA 469

```

Search completed: June 7, 2006, 05:47:51
Job time : 189 secs